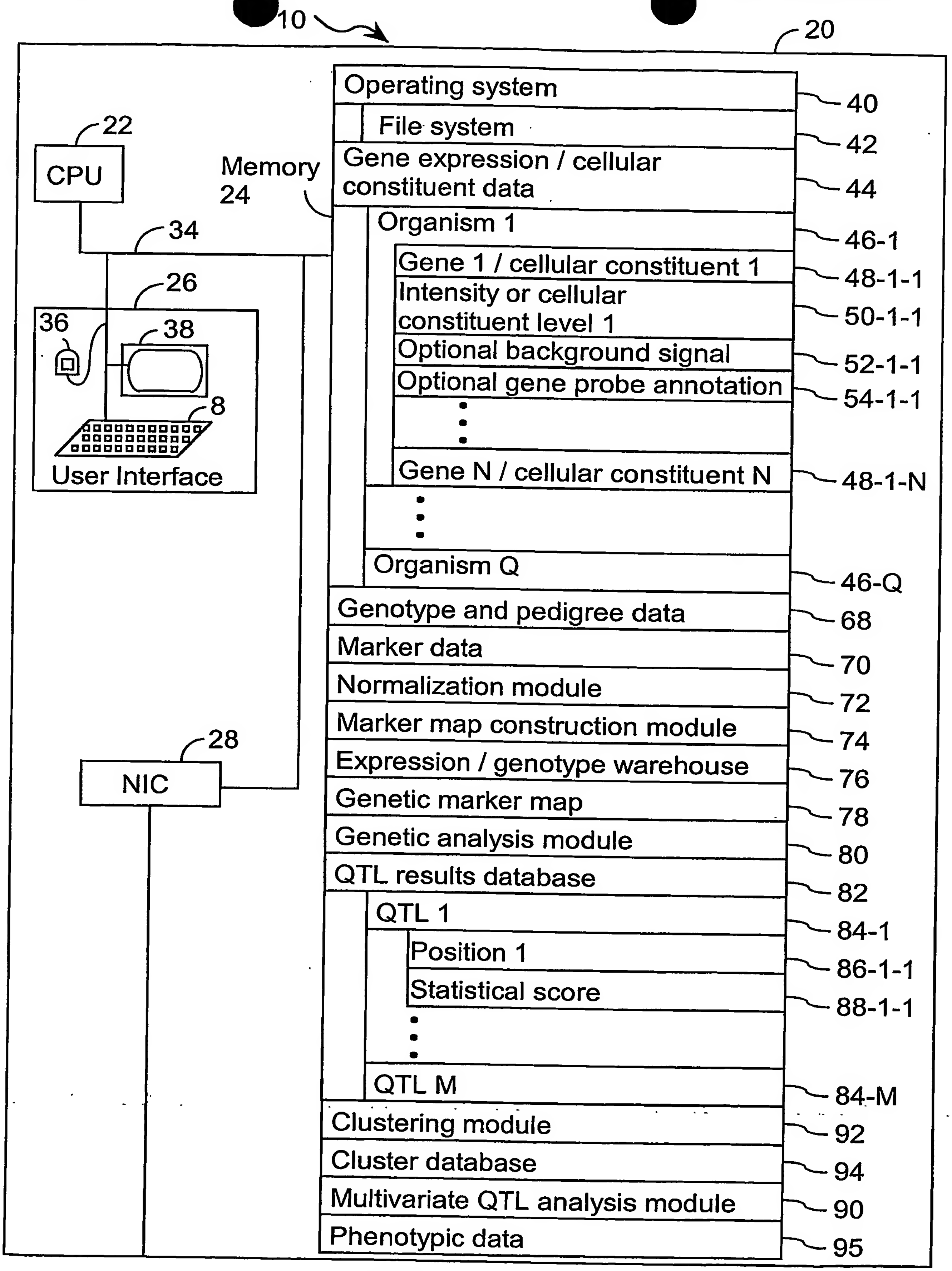

10
20
22
CPU
Memory
24
34
26
36
38
8
User Interface
28
NIC
Operating system
40
File system
42
Gene expression / cellular constituent data
44
Organism 1
46-1
Gene 1 / cellular constituent 1
48-1-1
Intensity or cellular constituent level 1
50-1-1
Optional background signal
52-1-1
Optional gene probe annotation
54-1-1
Gene N / cellular constituent N
48-1-N
Organism Q
46-Q
Genotype and pedigree data
68
Marker data
70
Normalization module
72
Marker map construction module
74
Expression / genotype warehouse
76
Genetic marker map
78
Genetic analysis module
80
QTL results database
82
QTL 1
84-1
Position 1
86-1-1
Statistical score
88-1-1
QTL M
84-M
Clustering module
92
Cluster database
94
Multivariate QTL analysis module
90
Phenotypic data
95

FIG. 1

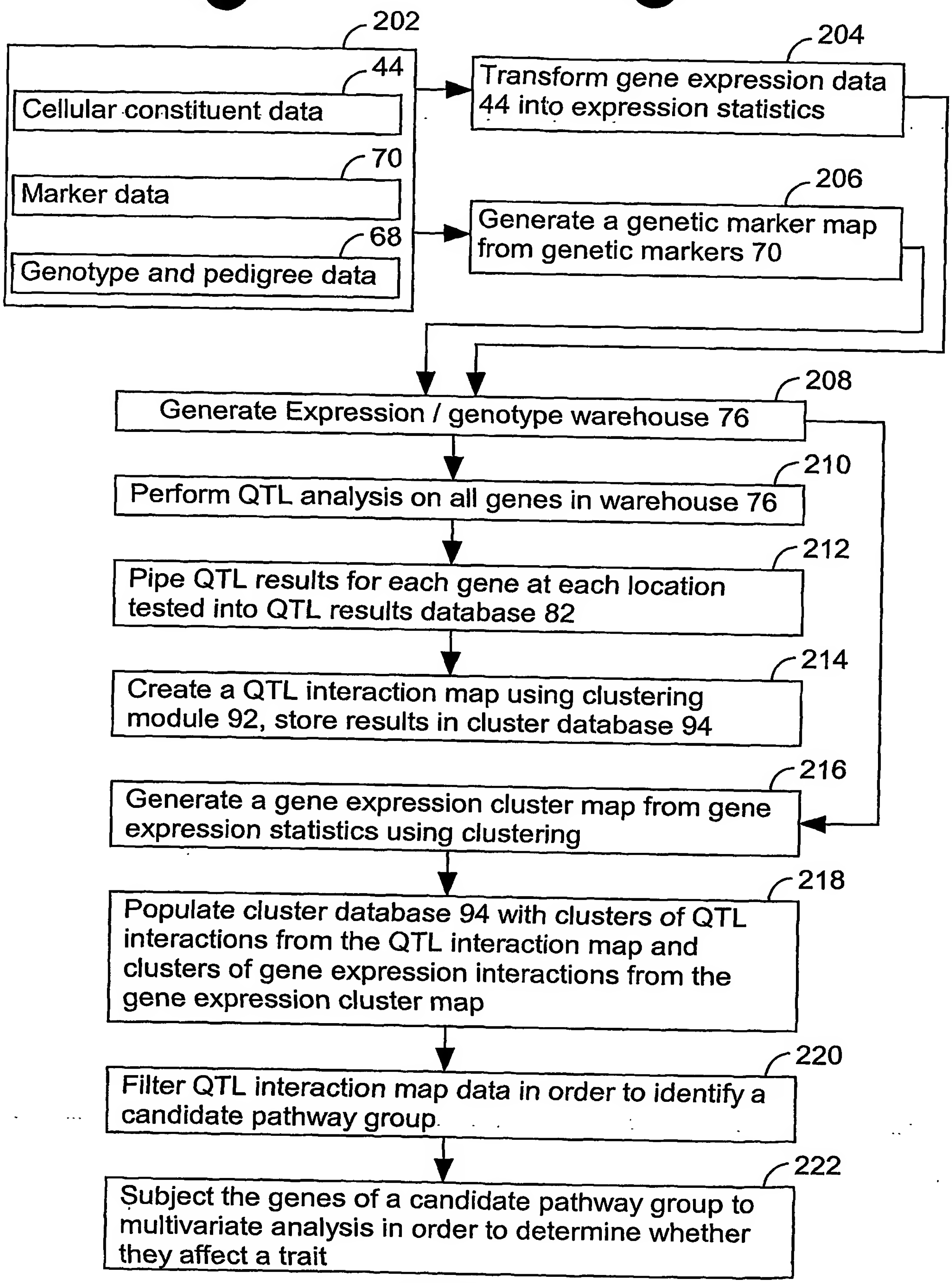

202
44
Cellular constituent data
70
Marker data
68
Genotype and pedigree data
204
Transform gene expression data 44 into expression statistics
206
Generate a genetic marker map from genetic markers 70
208
Generate Expression / genotype warehouse 76
210
Perform QTL analysis on all genes in warehouse 76
212
Pipe QTL results for each gene at each location tested into QTL results database 82
214
Create a QTL interaction map using clustering module 92, store results in cluster database 94
216
Generate a gene expression cluster map from gene expression statistics using clustering
218
Populate cluster database 94 with clusters of QTL interactions from the QTL interaction map and clusters of gene expression interactions from the gene expression cluster map
220
Filter QTL interaction map data in order to identify a candidate pathway group
222
Subject the genes of a candidate pathway group to multivariate analysis in order to determine whether they affect a trait

FIG. 2

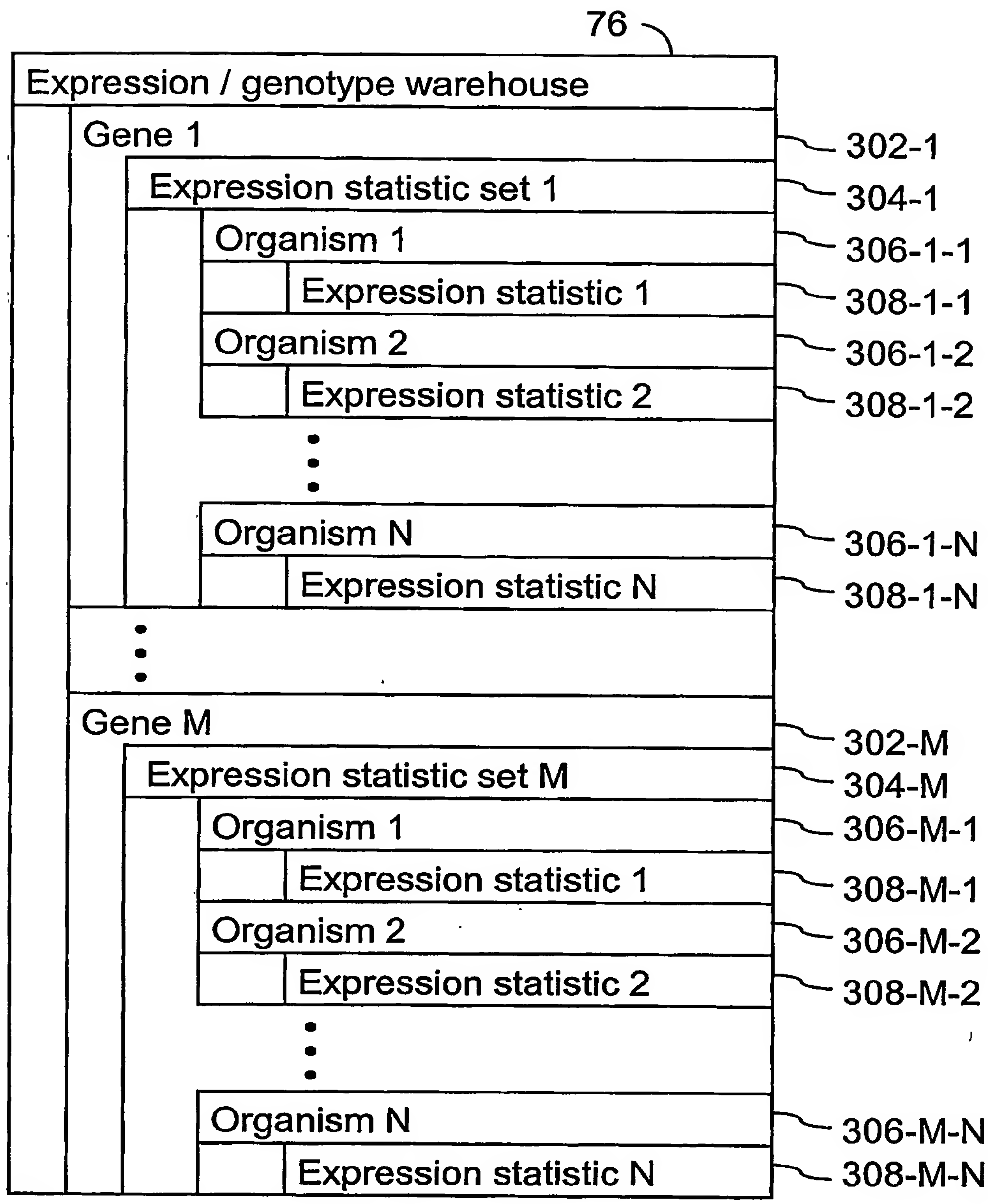

76
Expression / genotype warehouse
Gene 1
Expression statistic set 1
Organism 1
Expression statistic 1
Organism 2
Expression statistic 2
Organism N
Expression statistic N
Gene M
Expression statistic set M
Organism 1
Expression statistic 1
Organism 2
Expression statistic 2
Organism N
Expression statistic N
302-1
304-1
306-1-1
308-1-1
306-1-2
308-1-2
306-1-N
308-1-N
302-M
304-M
306-M-1
308-M-1
306-M-2
308-M-2
306-M-N
308-M-N

FIG. 3A

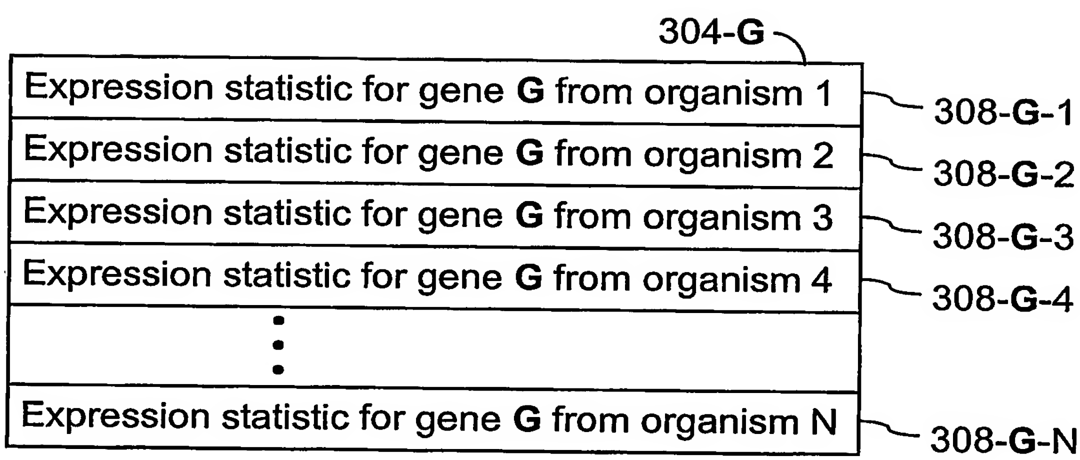

304-G
Expression statistic for gene G from organism 1
308-G-1
Expression statistic for gene G from organism 2
308-G-2
Expression statistic for gene G from organism 3
308-G-3
Expression statistic for gene G from organism 4
308-G-4
Expression statistic for gene G from organism N
308-G-N

FIG. 3B

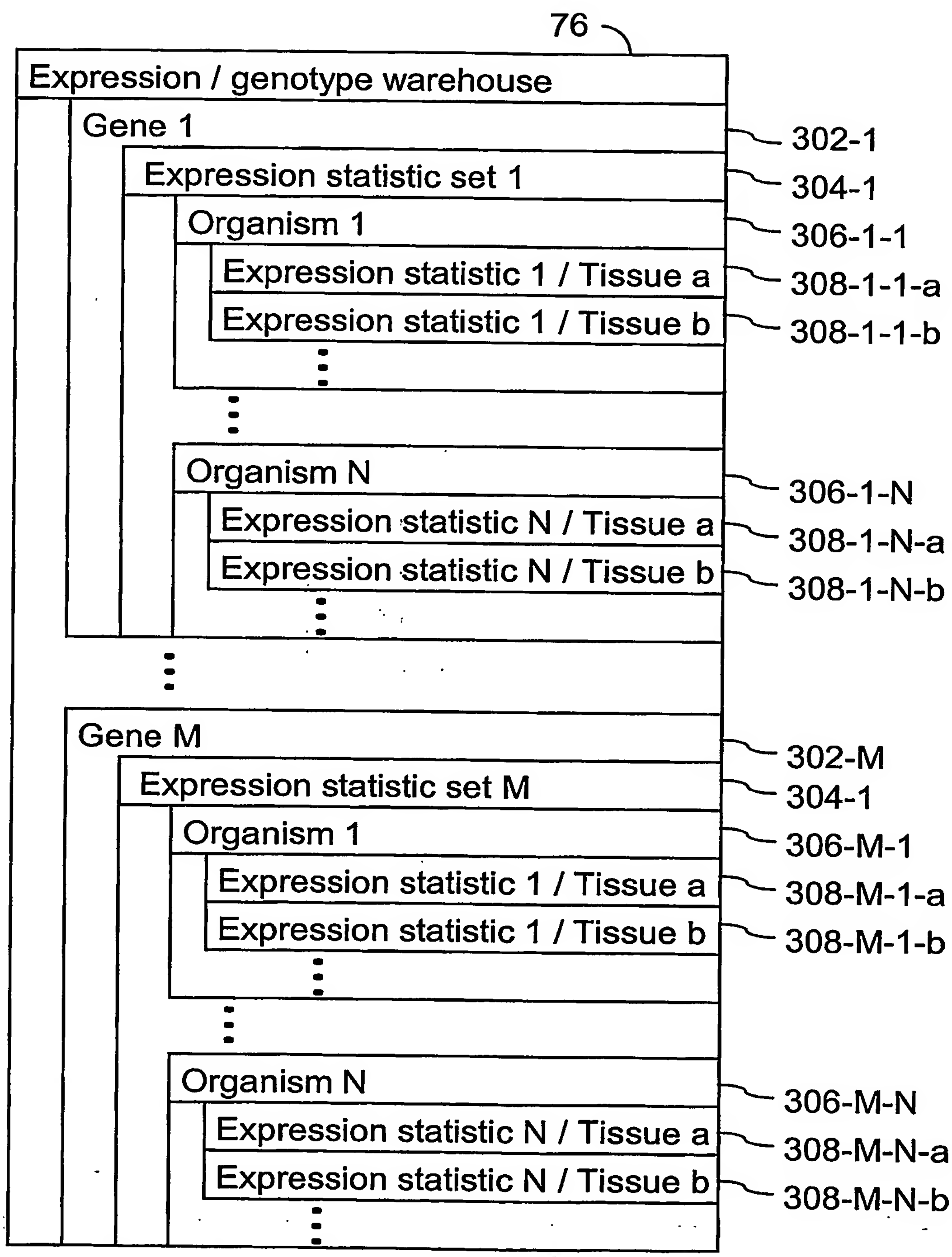

76
Expression / genotype warehouse
Gene 1
Expression statistic set 1
Organism 1
Expression statistic 1 / Tissue a
Expression statistic 1 / Tissue b
Organism N
Expression statistic N / Tissue a
Expression statistic N / Tissue b
Gene M
Expression statistic set M
Organism 1
Expression statistic 1 / Tissue a
Expression statistic 1 / Tissue b
Organism N
Expression statistic N / Tissue a
Expression statistic N / Tissue b
302-1
304-1
306-1-1
308-1-1-a
308-1-1-b
306-1-N
308-1-N-a
308-1-N-b
302-M
304-1
306-M-1
308-M-1-a
308-M-1-b
306-M-N
308-M-N-a
308-M-N-b

FIG. 3C

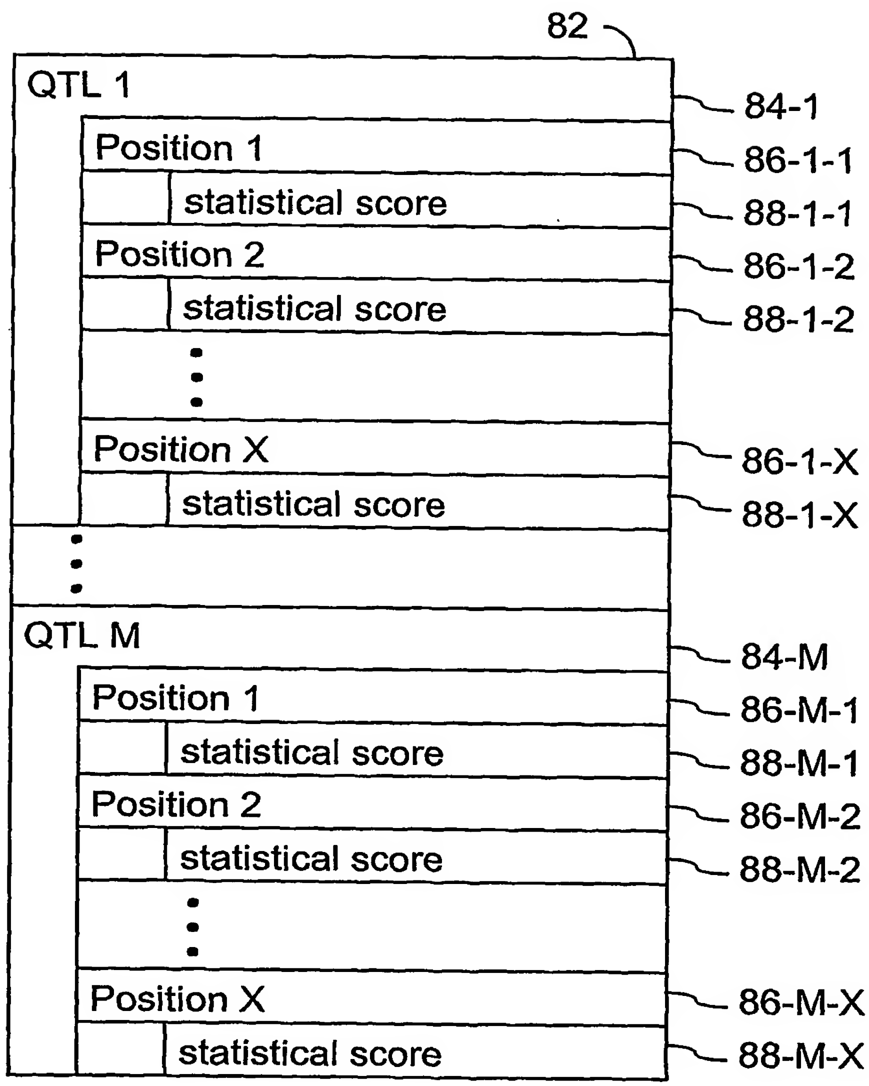

82
QTL 1
84-1
Position 1
86-1-1
statistical score
88-1-1
Position 2
86-1-2
statistical score
88-1-2
Position X
86-1-X
statistical score
88-1-X
QTL M
84-M
Position 1
86-M-1
statistical score
88-M-1
Position 2
86-M-2
statistical score
88-M-2
Position X
86-M-X
statistical score
88-M-X

FIG. 4

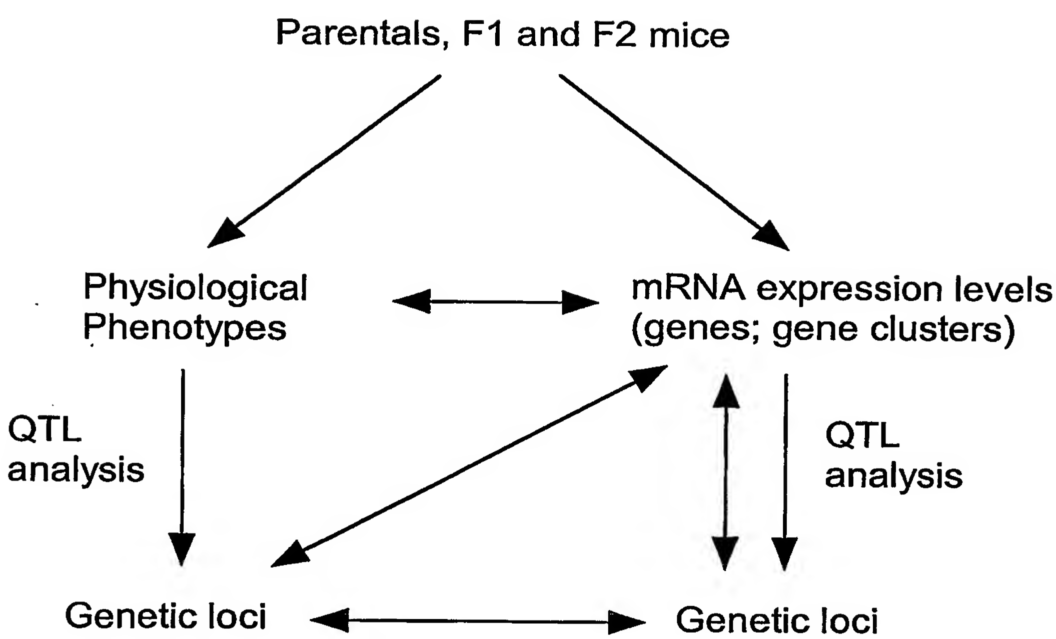

Parentals, F1 and F2 mice
Physiological Phenotypes
mRNA expression levels (genes; gene clusters)
QTL analysis
QTL analysis
Genetic loci
Genetic loci

FIG. 5

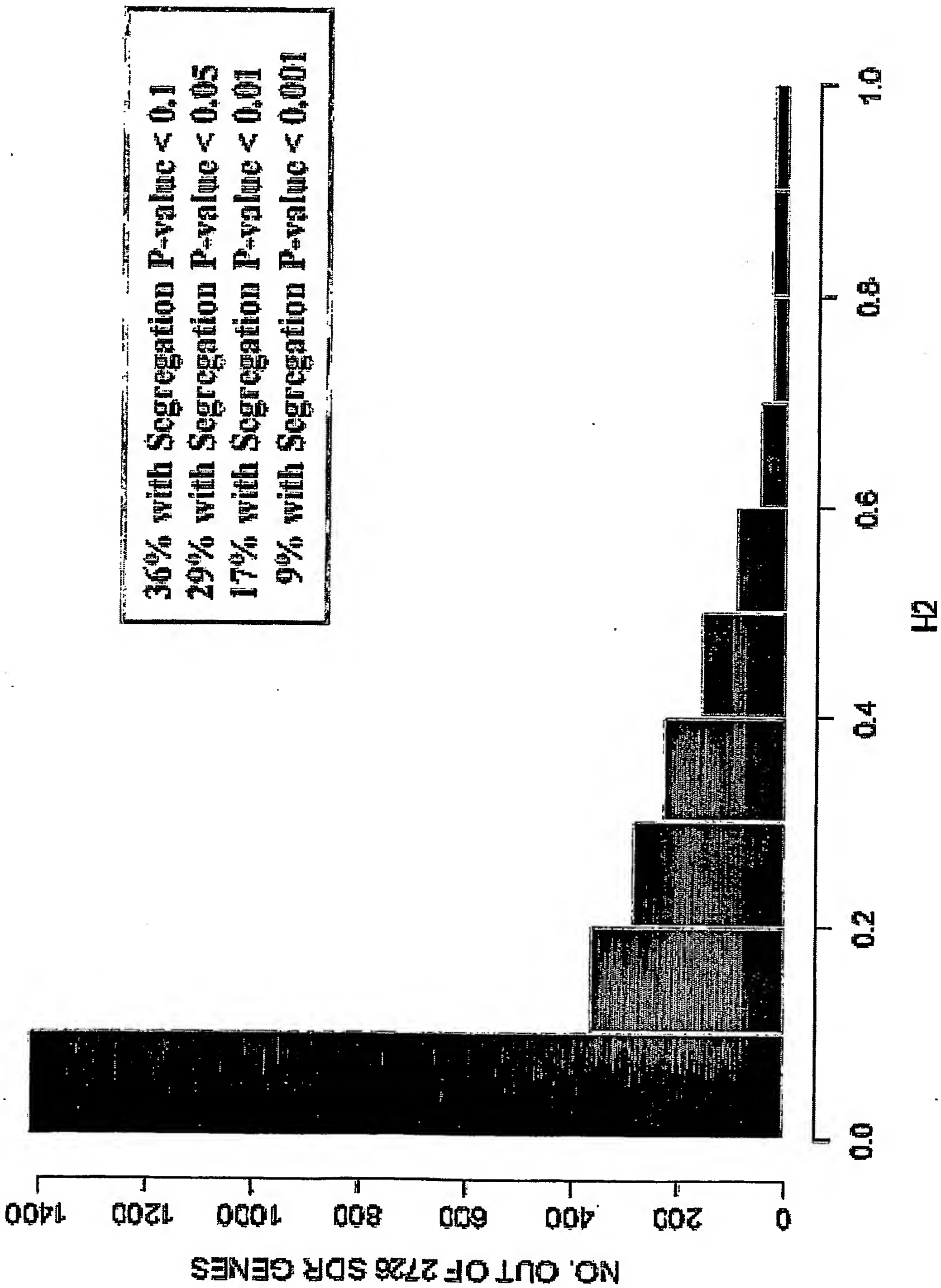

36% with Segregation P-value < 0.1
29% with Segregation P-value < 0.05
17% with Segregation P-value < 0.01
9% with Segregation P-value < 0.001
H2
NO. OUT OF 2726 SDR GENES
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
1200
1400

FIG. 6

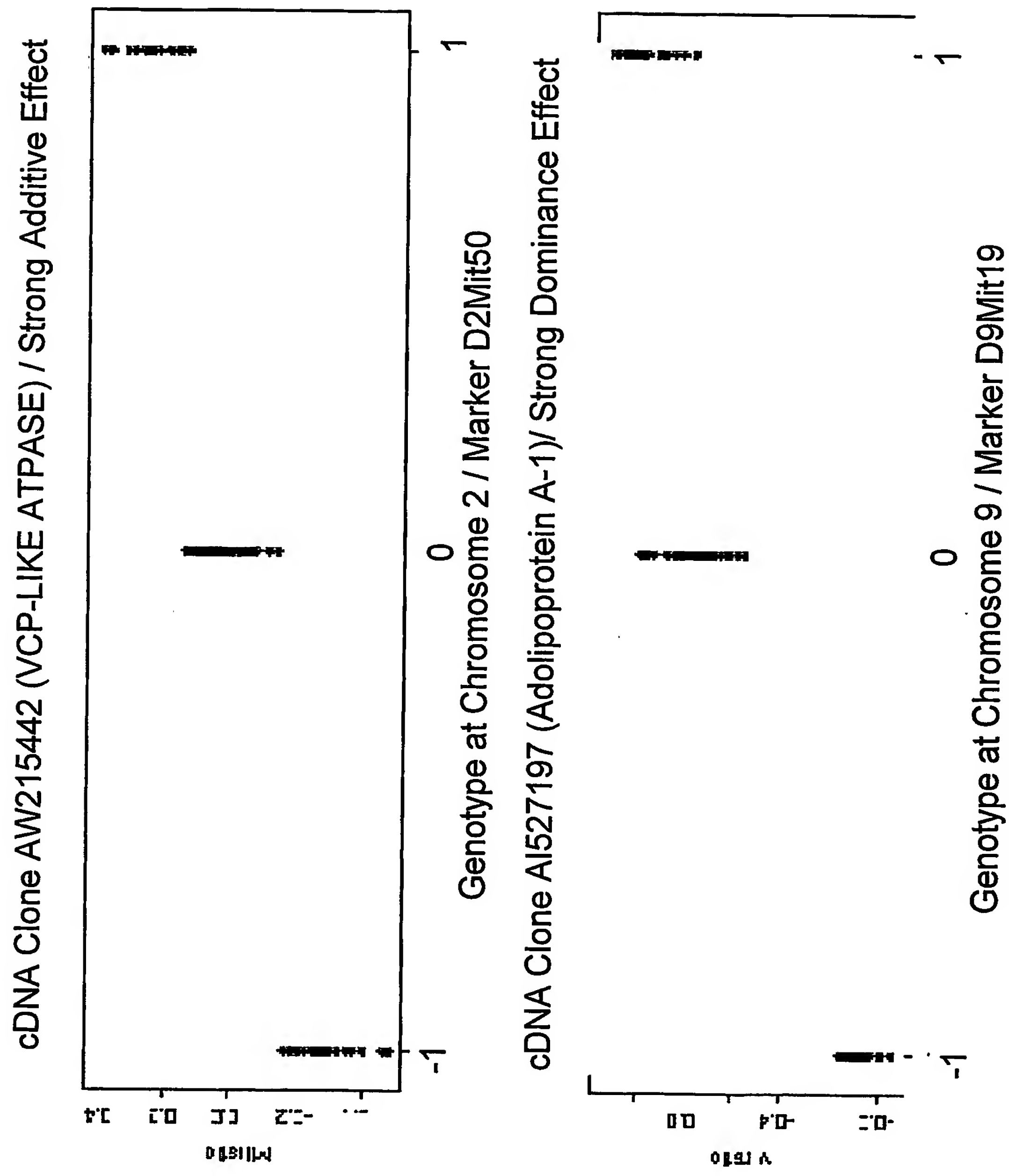

cDNA Clone AW215442 (VCP-LIKE ATPASE) / Strong Additive Effect
-1
0
1
Genotype at Chromosome 2 / Marker D2Mit50
cDNA Clone AI527197 (Adolipoprotein A-1)/ Strong Dominance Effect
-1
0
1
Genotype at Chromosome 9 / Marker D9Mit19

FIG. 7

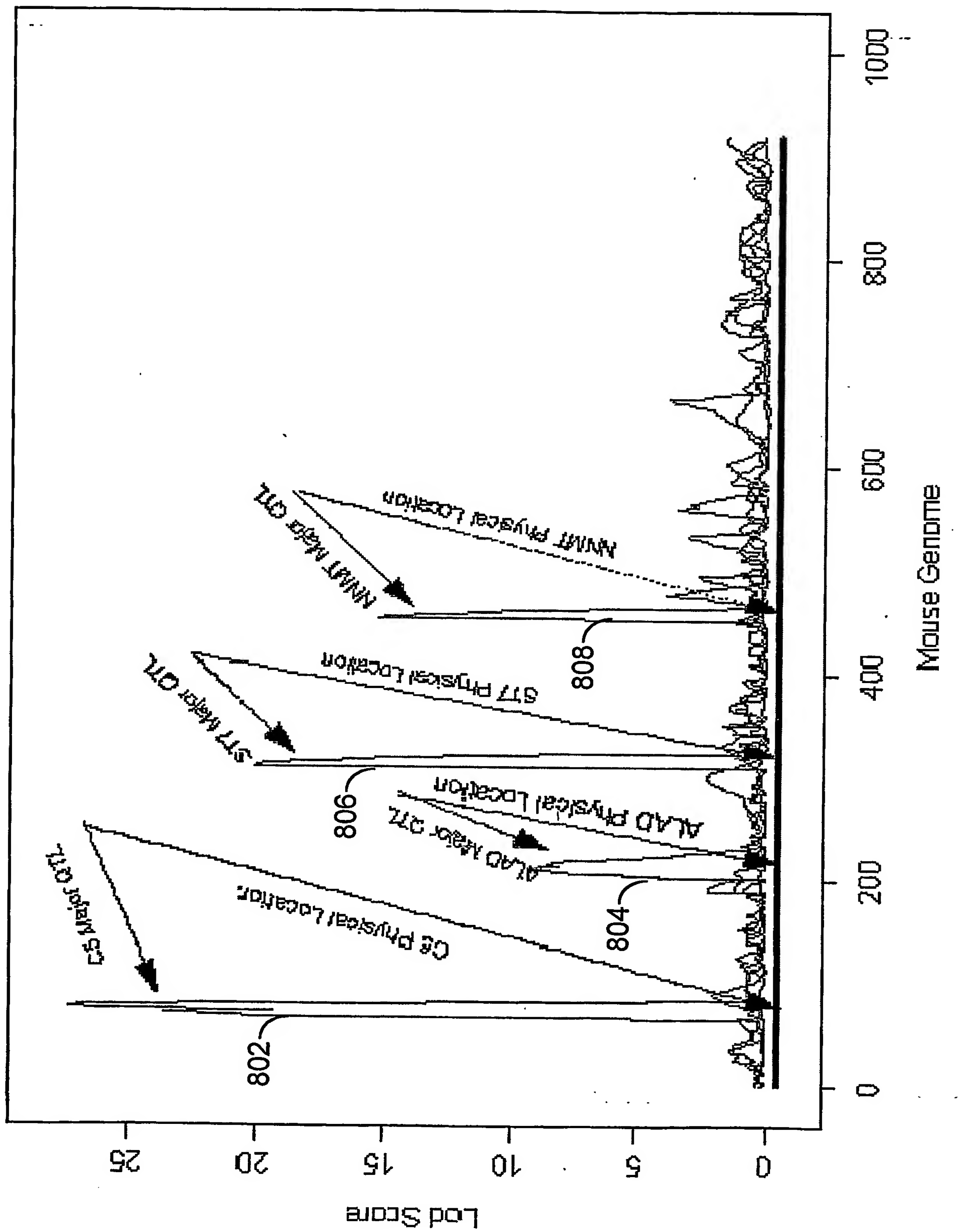

C5 Major QTL
C5 Physical Location
ALAD Major QTL
ALAD Physical Location
ST7 Major QTL
ST7 Physical Location
NNMT Major QTL
NNMT Physical Location
802
804
806
808
Mouse Genome
Lod Score
0
200
400
600
800
1000
0
5
10
15
20
25

FIG. 8

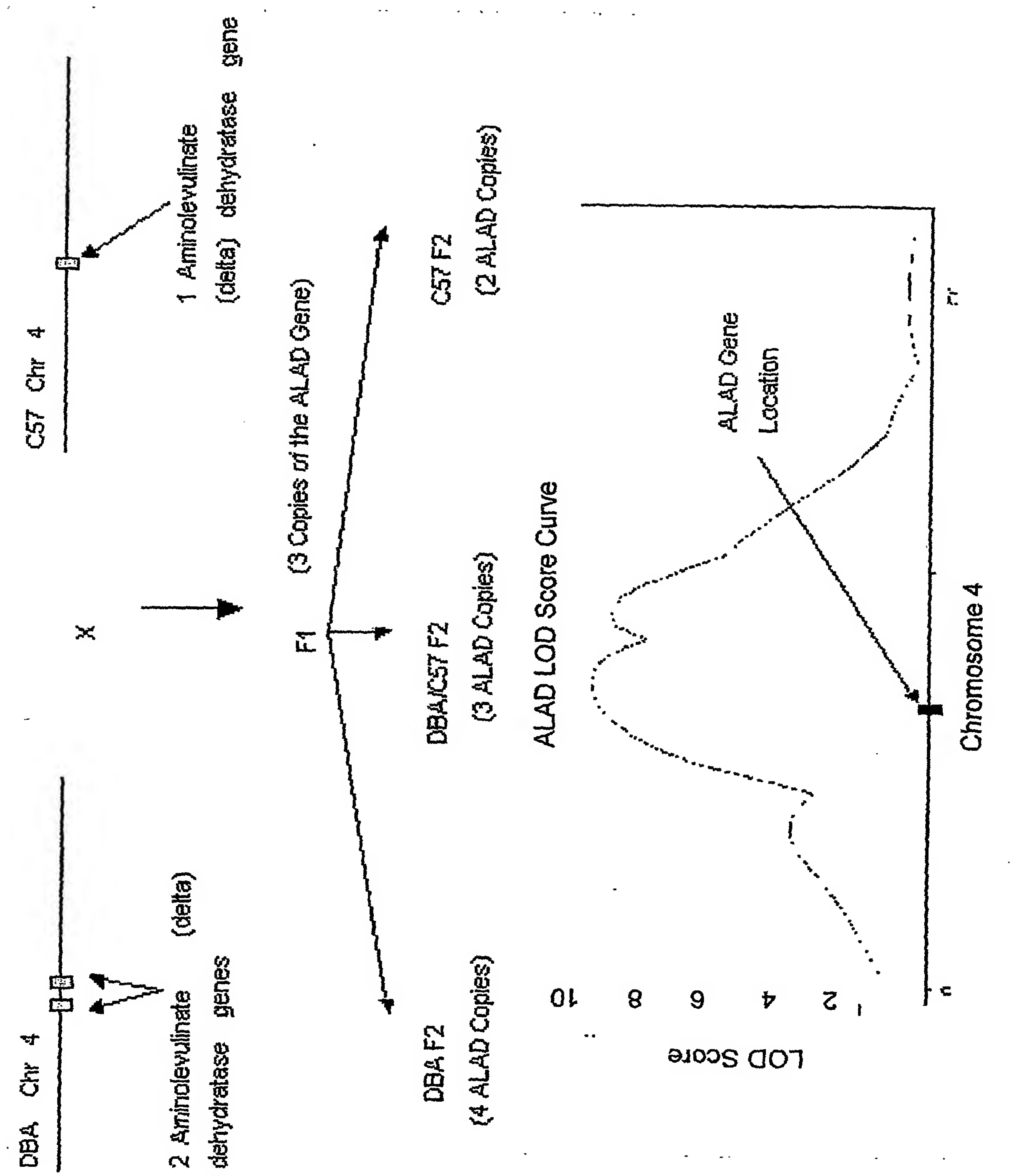

DBA Chr 4
2 Aminolevulinate dehydratase genes
(delta)
C57 Chr 4
1 Aminolevulinate (delta) dehydratase gene
X
F1
(3 Copies of the ALAD Gene)
DBA F2
(4 ALAD Copies)
DBA/C57 F2
(3 ALAD Copies)
C57 F2
(2 ALAD Copies)
ALAD LOD Score Curve
LOD Score
10
8
6
4
2
ALAD Gene Location
Chromosome 4

FIG. 9

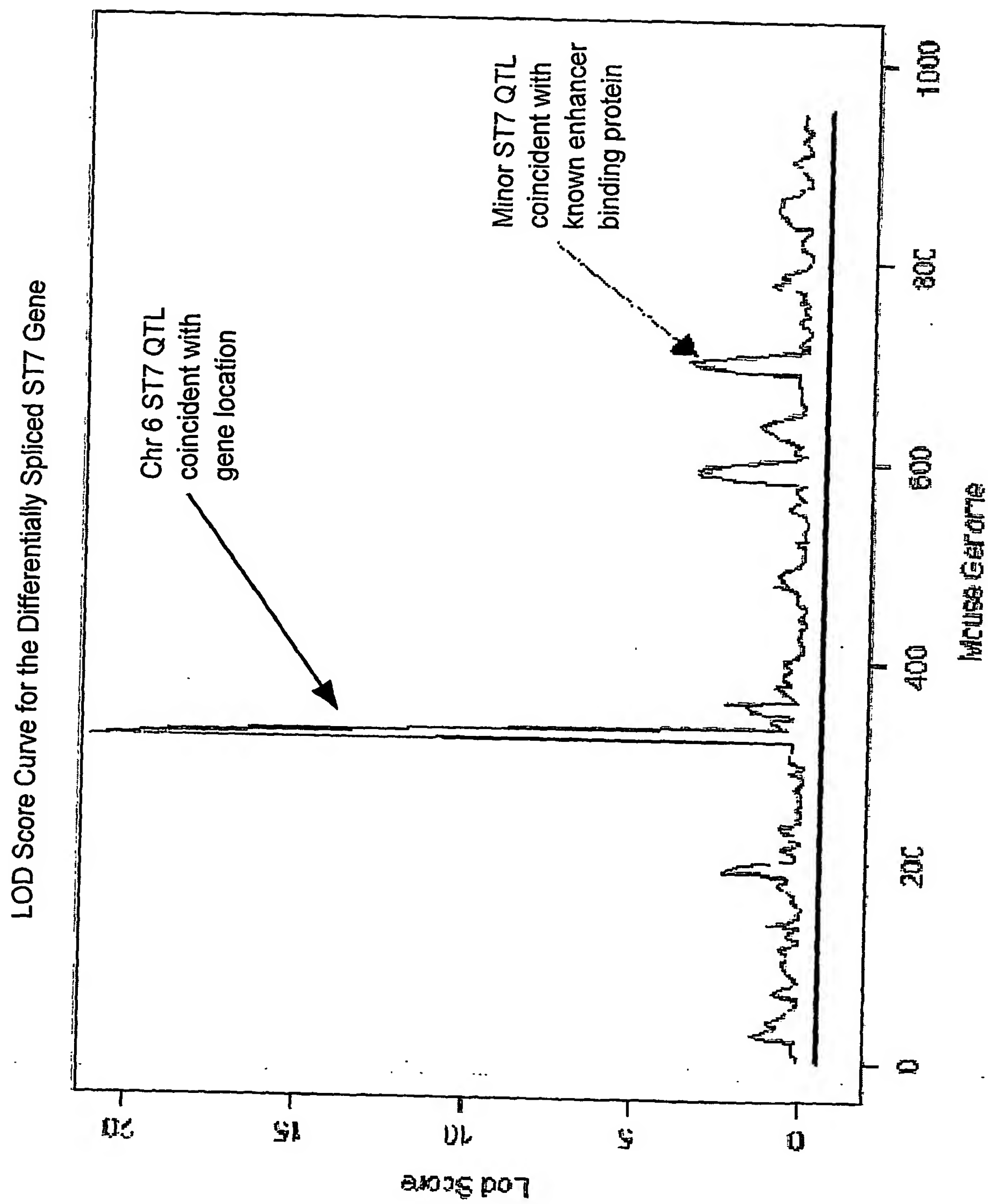
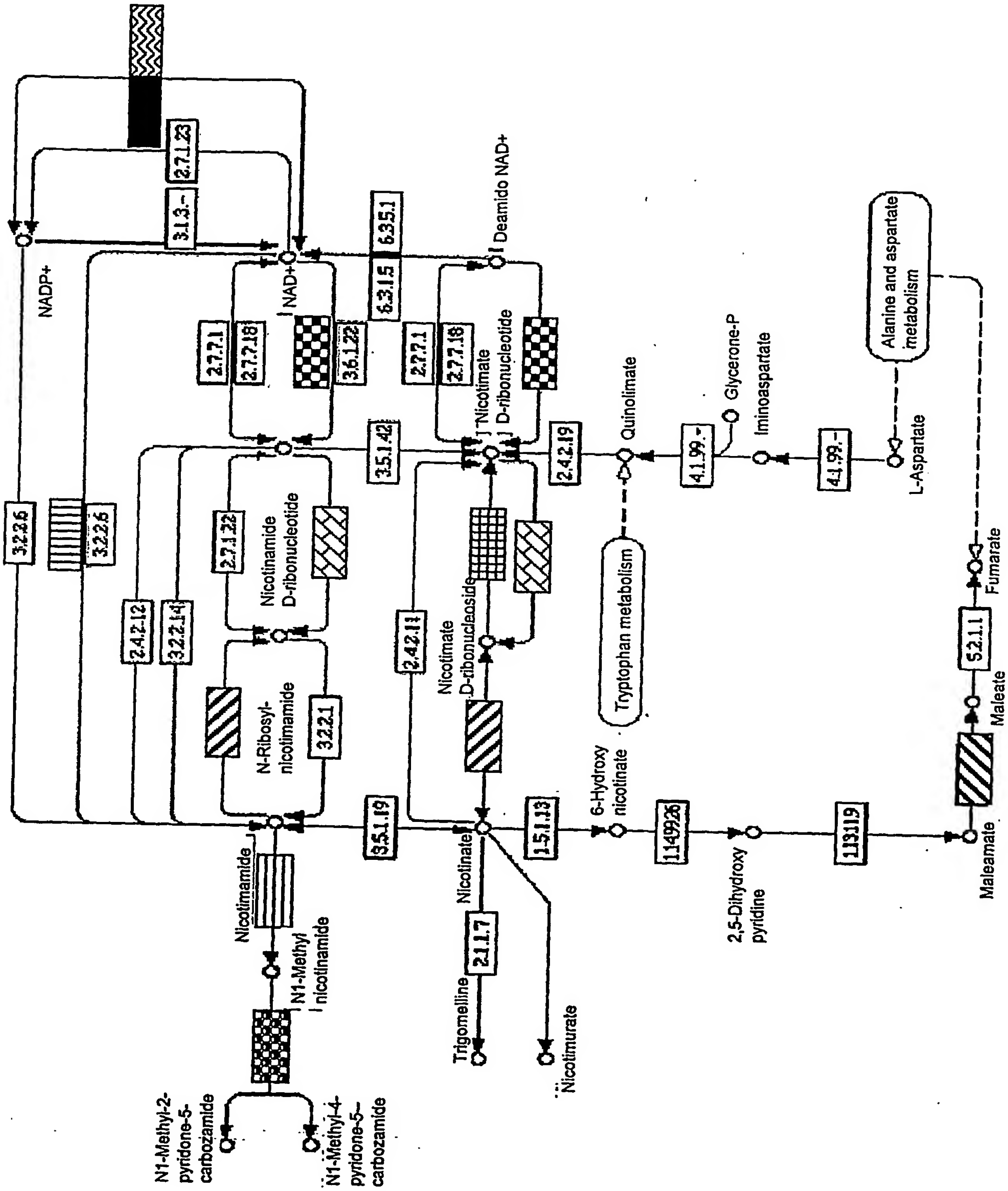
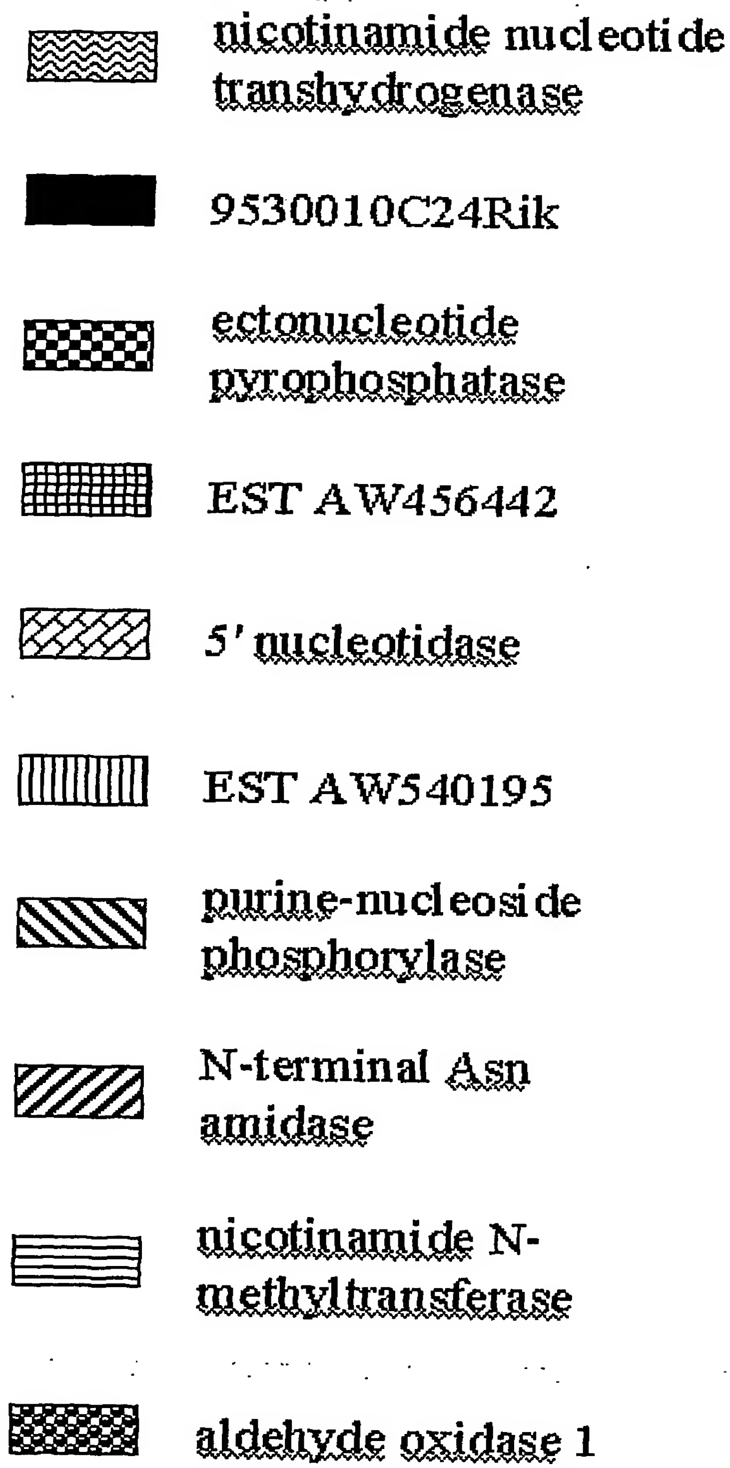

LOD Score Curve for the Differentially Spliced ST7 Gene
Chr 6 ST7 QTL coincident with gene location
Minor ST7 QTL coincident with known enhancer binding protein
Lod Score
0
5
10
15
20
Mouse Genome
0
200
400
600
800
1000

FIG. 10




**(PRIOR ART)**
**FIG. 11**




**(PRIOR ART)**
**FIG. 12**

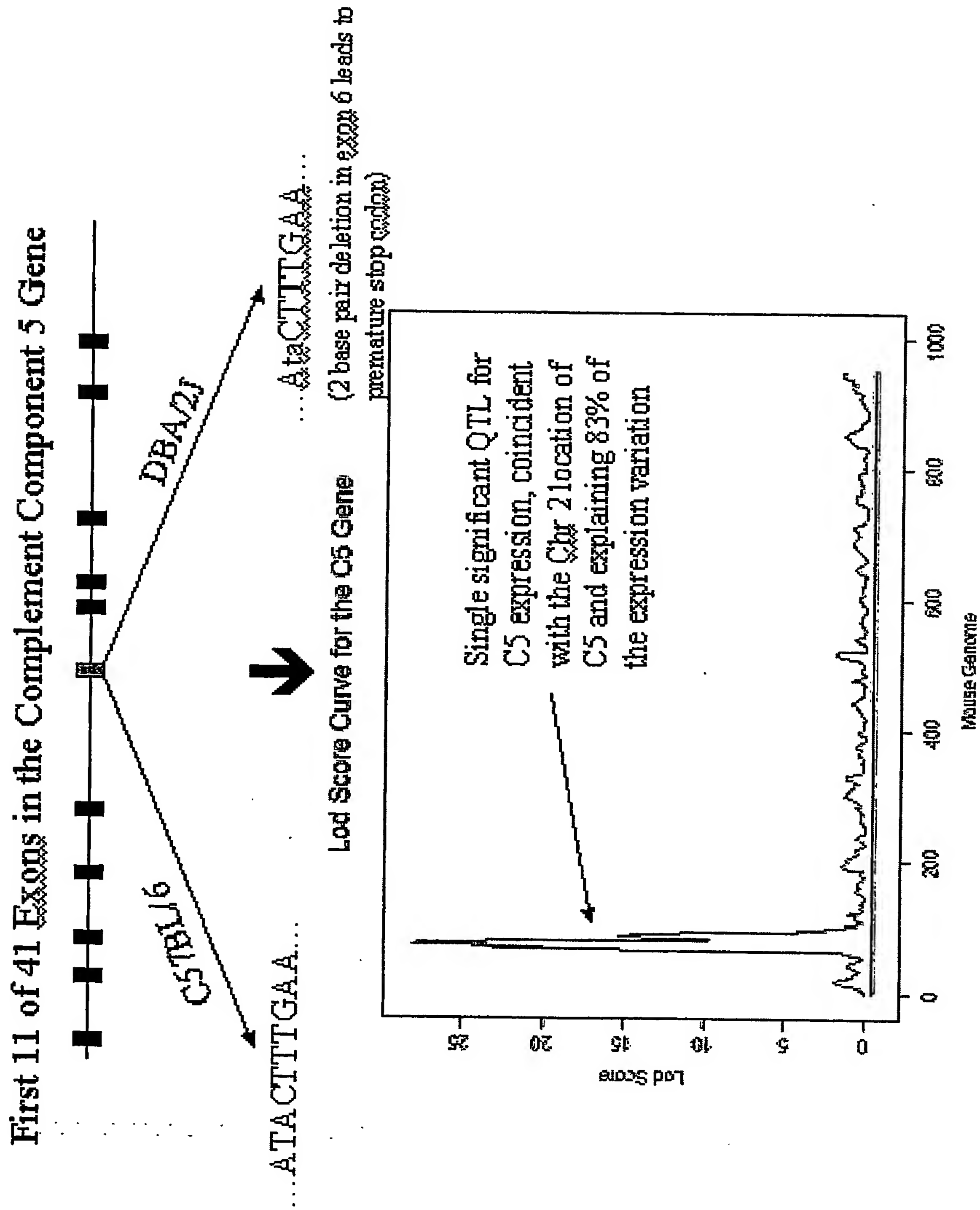

First 11 of 41 Exons in the Complement Component 5 Gene
C57BL/6
DBA/2J
....ATACTTTGAA....
....AtaCTTTGAA....
(2 base pair deletion in exon 6 leads to premature stop codon)
Lod Score Curve for the C5 Gene
Single significant QTL for C5 expression, coincident with the Chr 2 location of C5 and explaining 83% of the expression variation
Lod Score
0
5
10
15
20
25
Mouse Genome
0
200
400
600
800
1000

**FIG. 13**

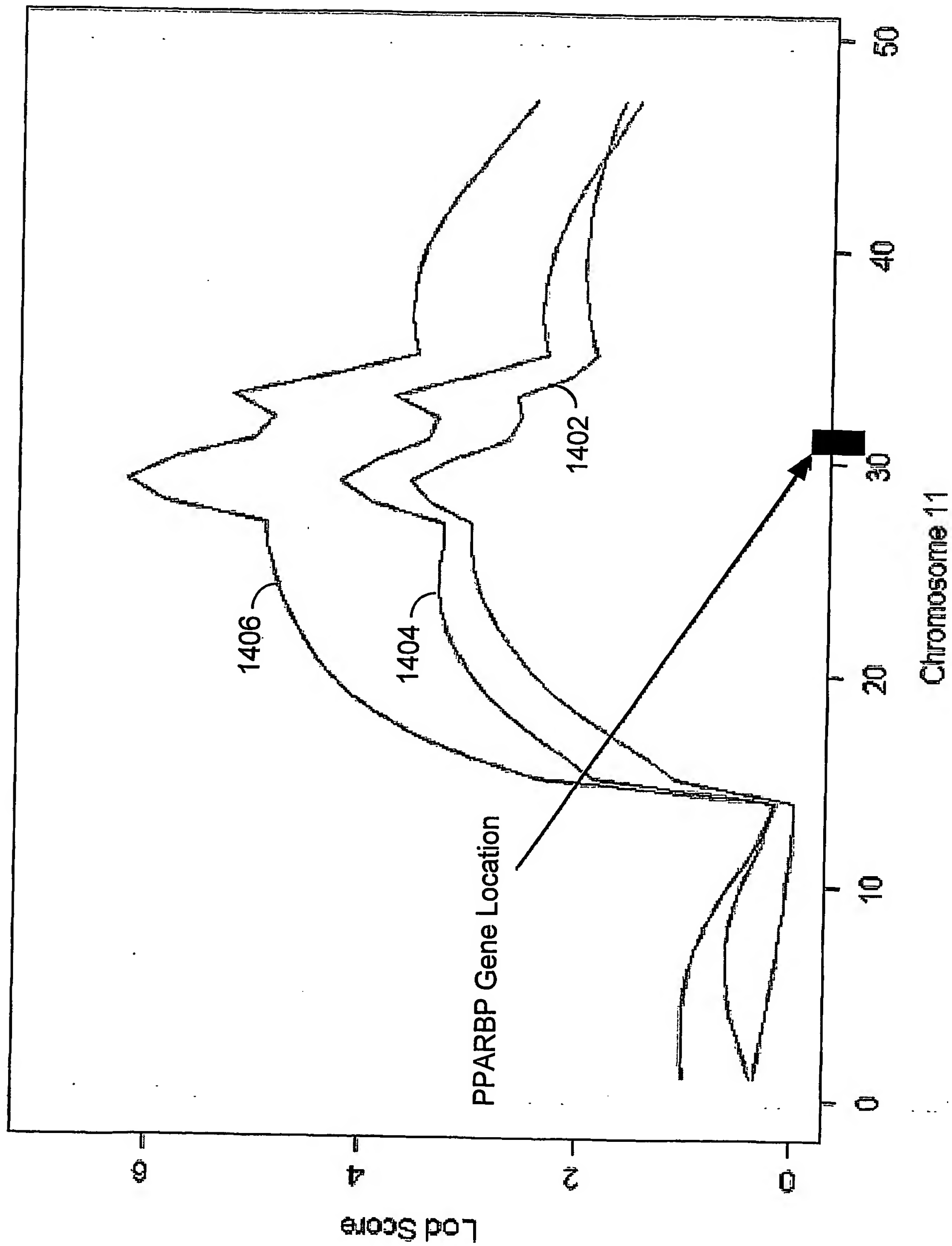

PPARBP Gene Location
1402
1404
1406
Lod Score
0
2
4
6
Chromosome 11
0
10
20
30
40
50

**FIG. 14**

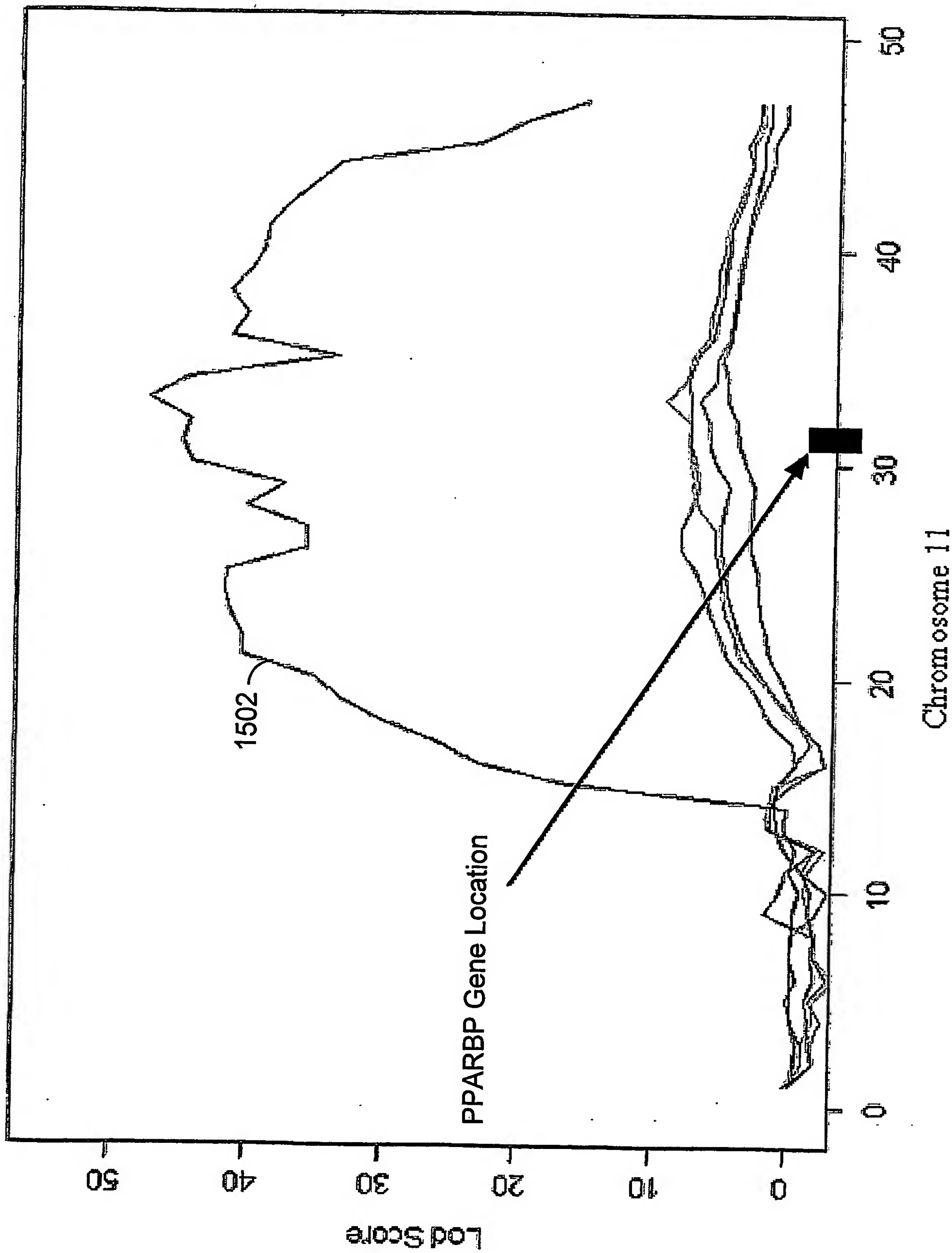

1502
PPARBP Gene Location
Chromosome 11
0
10
20
30
40
50
Lod Score
0
10
20
30
40
50

FIG. 15

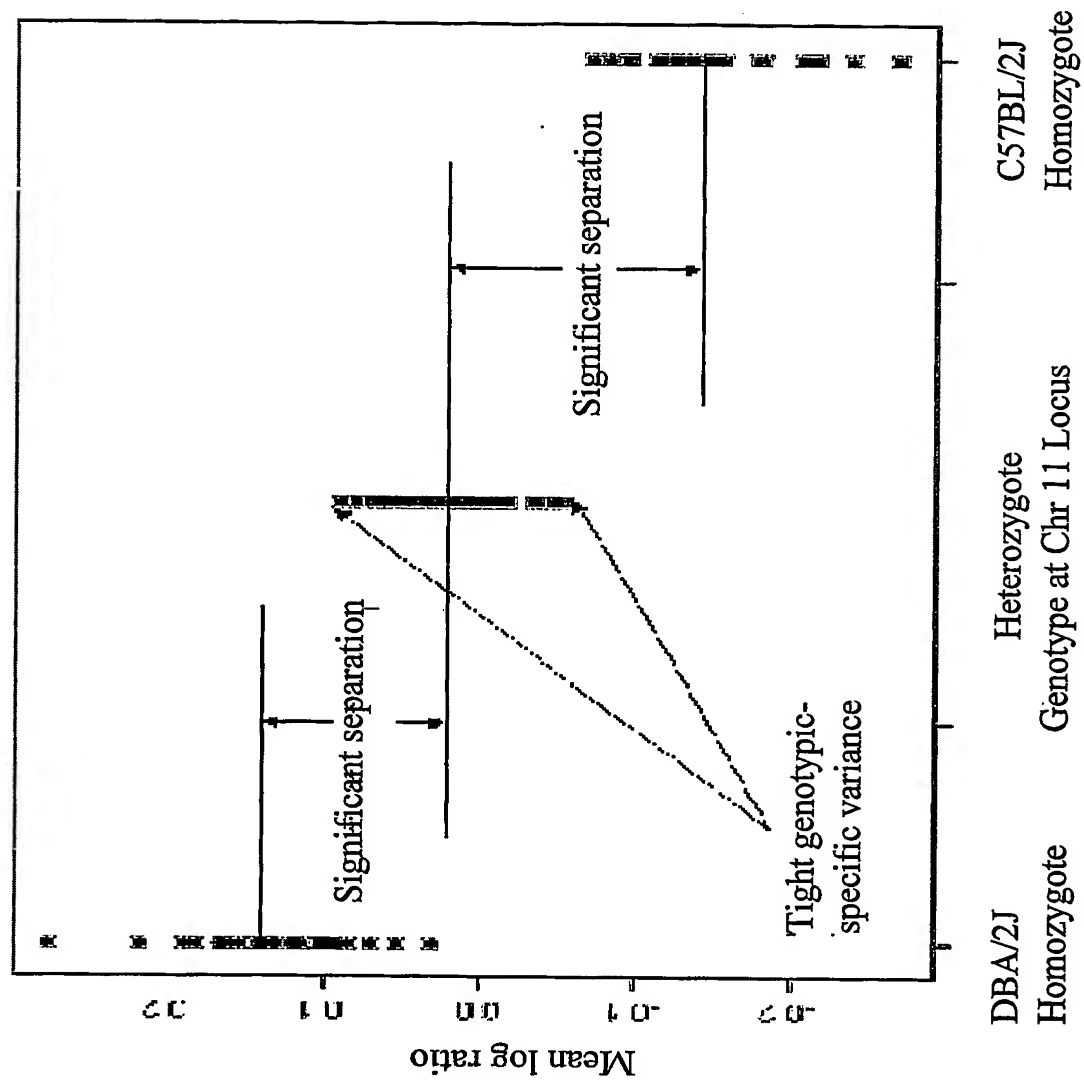

Significant separation
Significant separation
Tight genotypic-specific variance
Mean log ratio
DBA/2J Homozygote
Heterozygote
Genotype at Chr 11 Locus
C57BL/2J Homozygote

**FIG. 16**

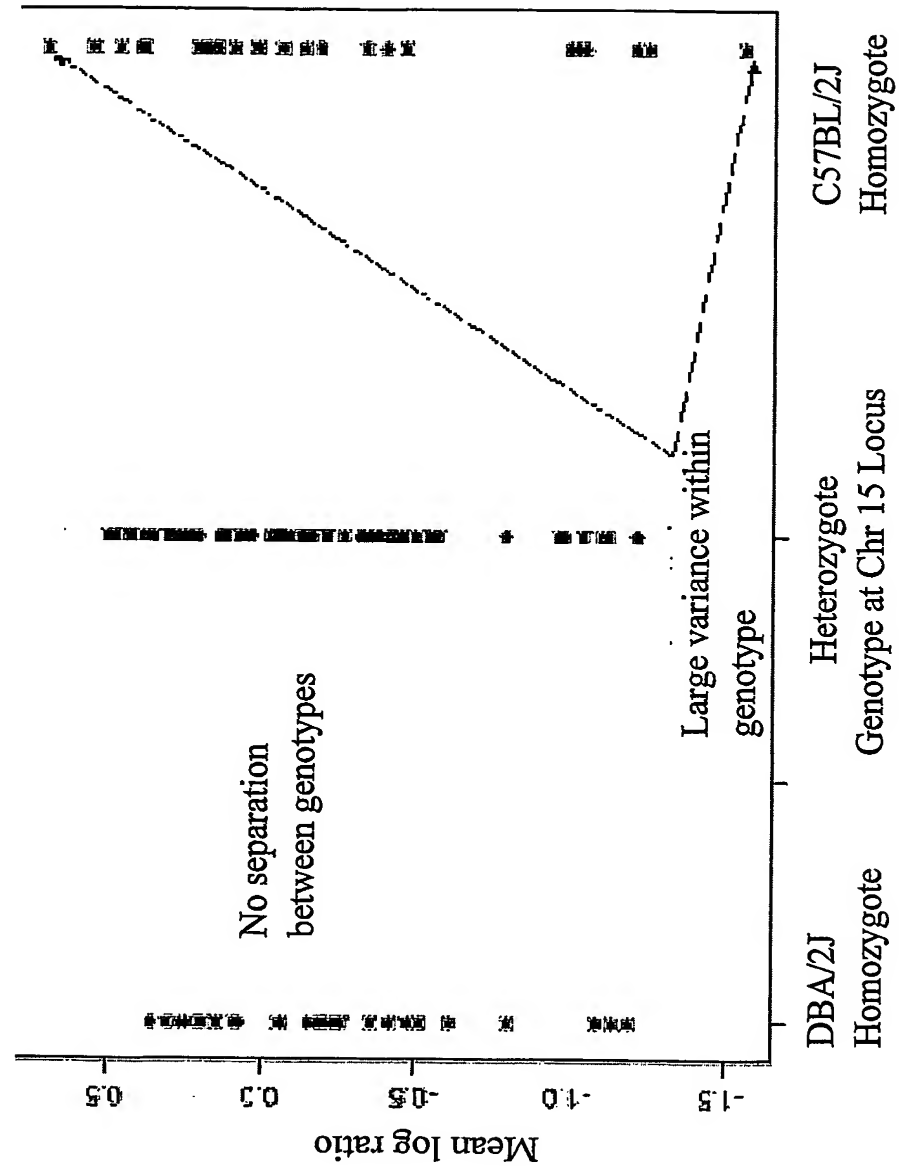

No separation
between genotypes
Large variance within
genotype
Mean log ratio
0.5
0.0
-0.5
-1.0
-1.5
DBA/2J
Homozygote
Heterozygote
C57BL/2J
Homozygote
Genotype at Chr 15 Locus

**FIG. 17**

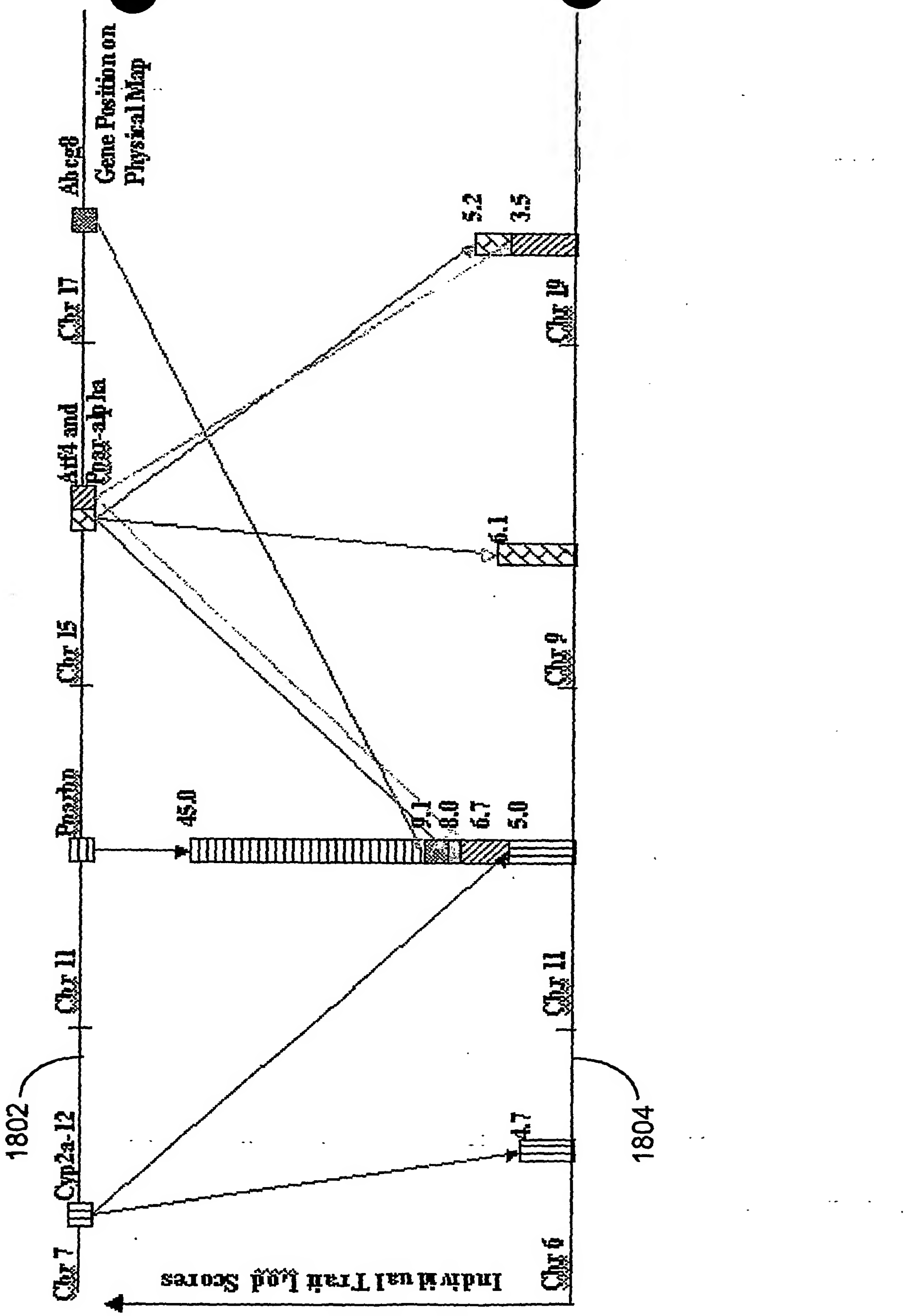

Chr 7
Cyp2a-12
Chr 11
Pparbp
Chr 15
Atf4 and Ppar-alpha
Chr 17
Abcg8
Gene Position on Physical Map
1802
45.0
9.1
8.0
6.7
5.0
6.1
5.2
3.5
4.7
Chr 6
Chr 11
Chr 9
Chr 19
1804
Individual Trait Lod Scores

FIG. 18

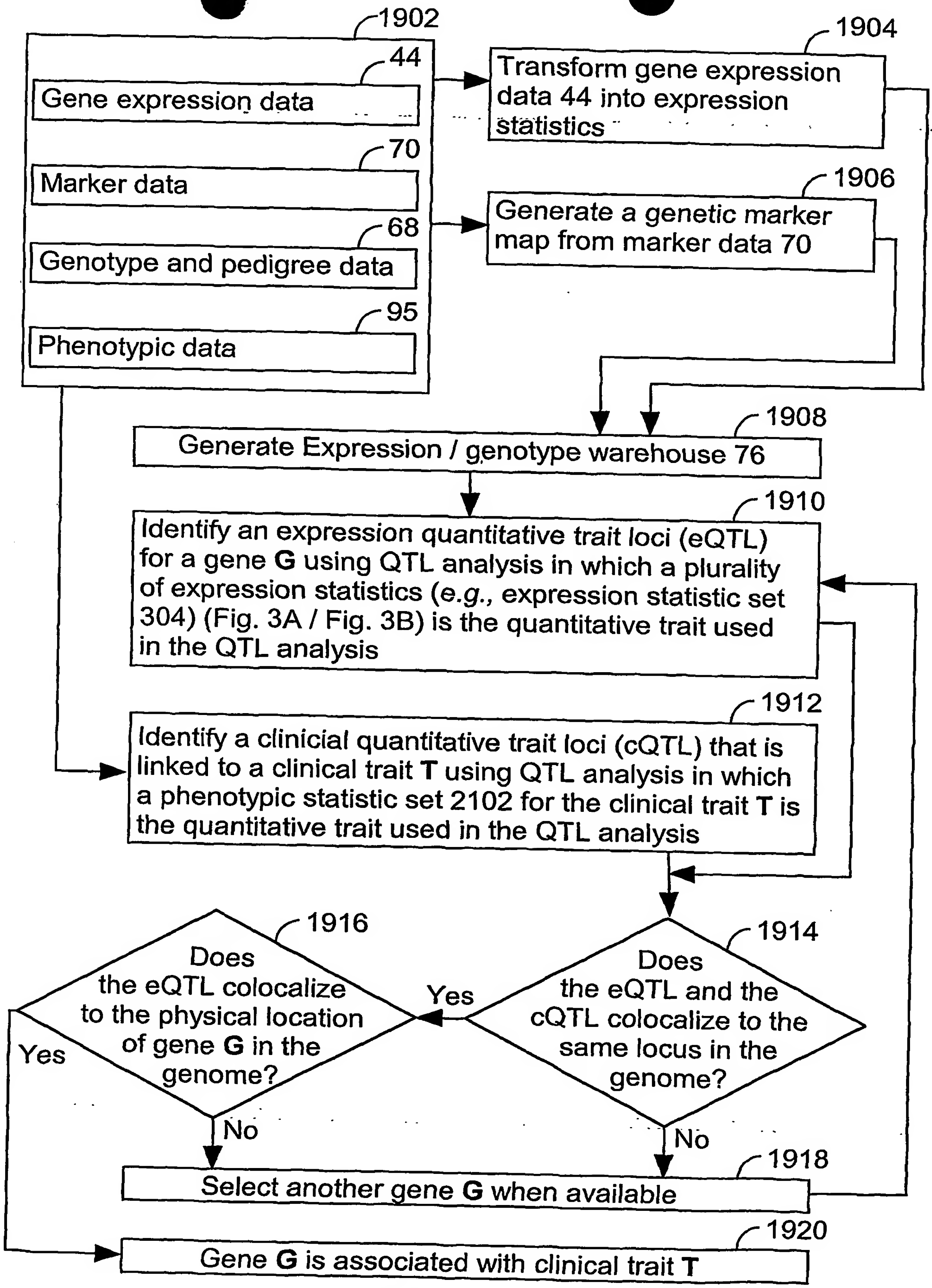

1902
44
Gene expression data
70
Marker data
68
Genotype and pedigree data
95
Phenotypic data
1904
Transform gene expression data 44 into expression statistics
1906
Generate a genetic marker map from marker data 70
1908
Generate Expression / genotype warehouse 76
1910
Identify an expression quantitative trait loci (eQTL) for a gene G using QTL analysis in which a plurality of expression statistics (e.g., expression statistic set 304) (Fig. 3A / Fig. 3B) is the quantitative trait used in the QTL analysis
1912
Identify a clinicial quantitative trait loci (cQTL) that is linked to a clinical trait T using QTL analysis in which a phenotypic statistic set 2102 for the clinical trait T is the quantitative trait used in the QTL analysis
1916
Does the eQTL colocalize to the physical location of gene G in the genome?
Yes
1914
Does the eQTL and the cQTL colocalize to the same locus in the genome?
Yes
No
No
1918
Select another gene G when available
1920
Gene G is associated with clinical trait T

FIG. 19

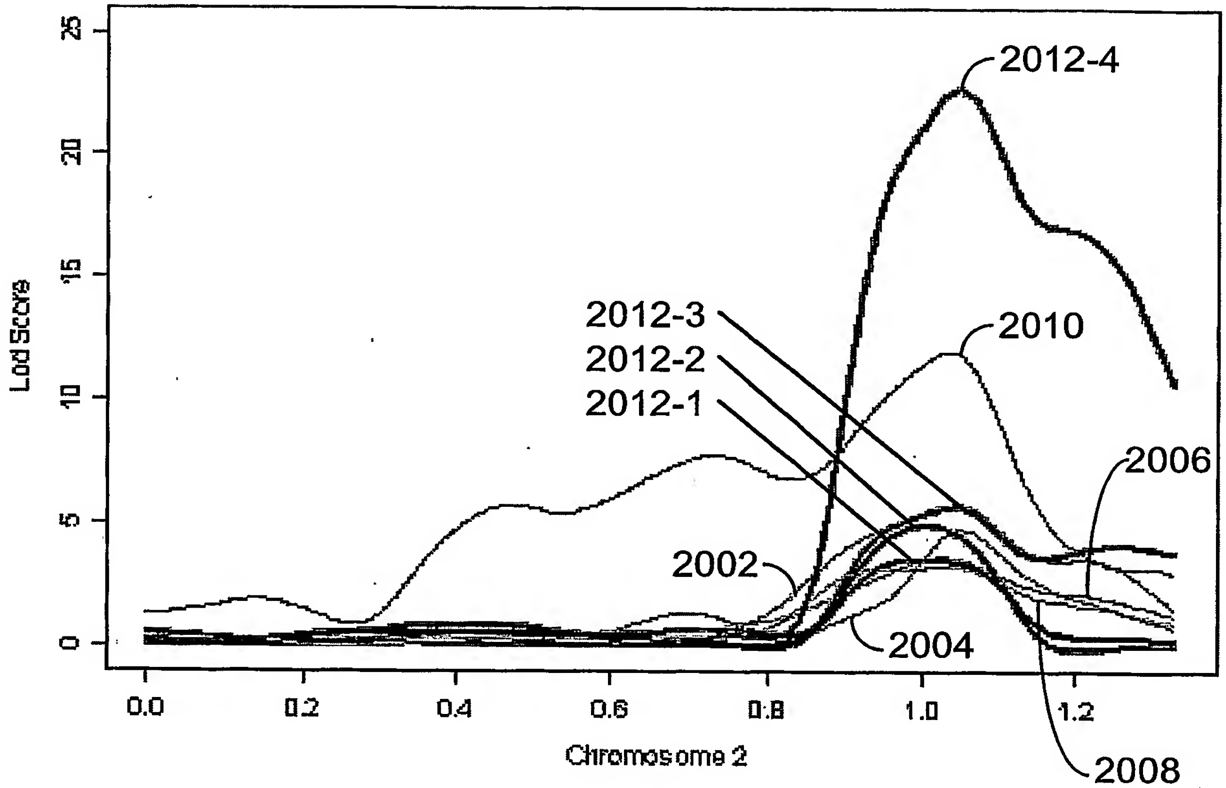

2012-4
2010
2012-3
2012-2
2012-1
2006
2002
2004
2008
Lod Score
25
20
15
10
5
0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Chromosome 2

**FIG. 20**

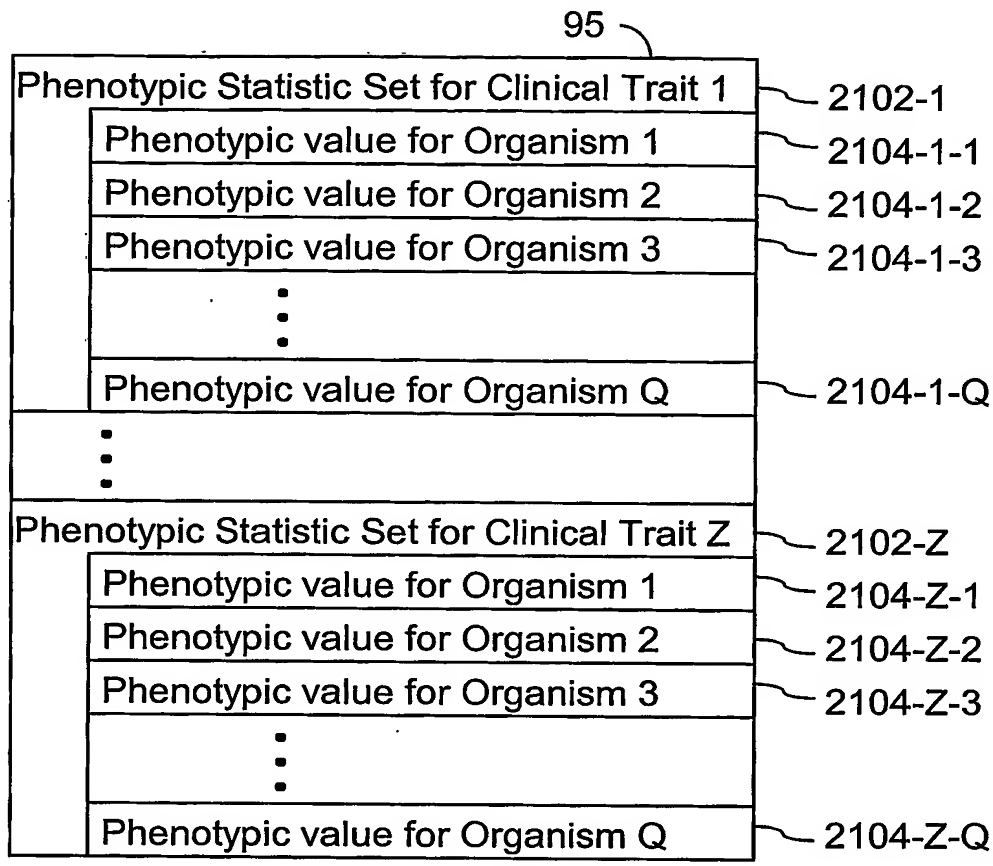

95
Phenotypic Statistic Set for Clinical Trait 1
Phenotypic value for Organism 1
Phenotypic value for Organism 2
Phenotypic value for Organism 3
Phenotypic value for Organism Q
Phenotypic Statistic Set for Clinical Trait Z
Phenotypic value for Organism 1
Phenotypic value for Organism 2
Phenotypic value for Organism 3
Phenotypic value for Organism Q
2102-1
2104-1-1
2104-1-2
2104-1-3
2104-1-Q
2102-Z
2104-Z-1
2104-Z-2
2104-Z-3
2104-Z-Q

**FIG. 21**

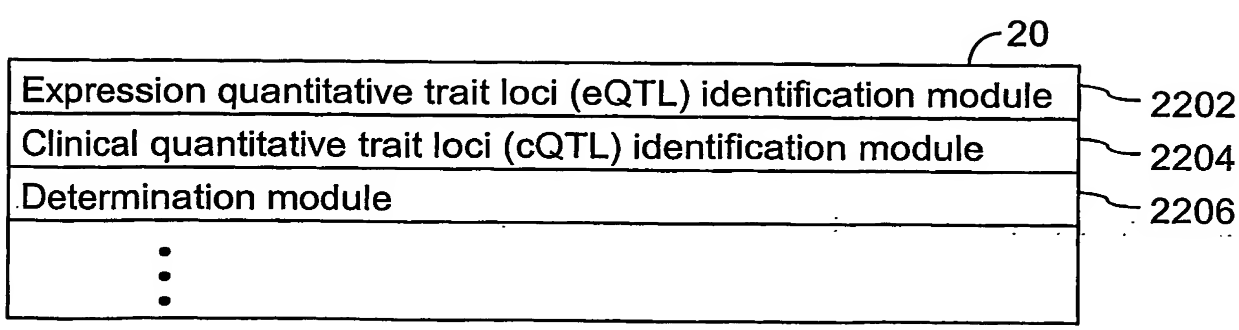

20
Expression quantitative trait loci (eQTL) identification module
Clinical quantitative trait loci (cQTL) identification module
Determination module
2202
2204
2206

**FIG. 22**

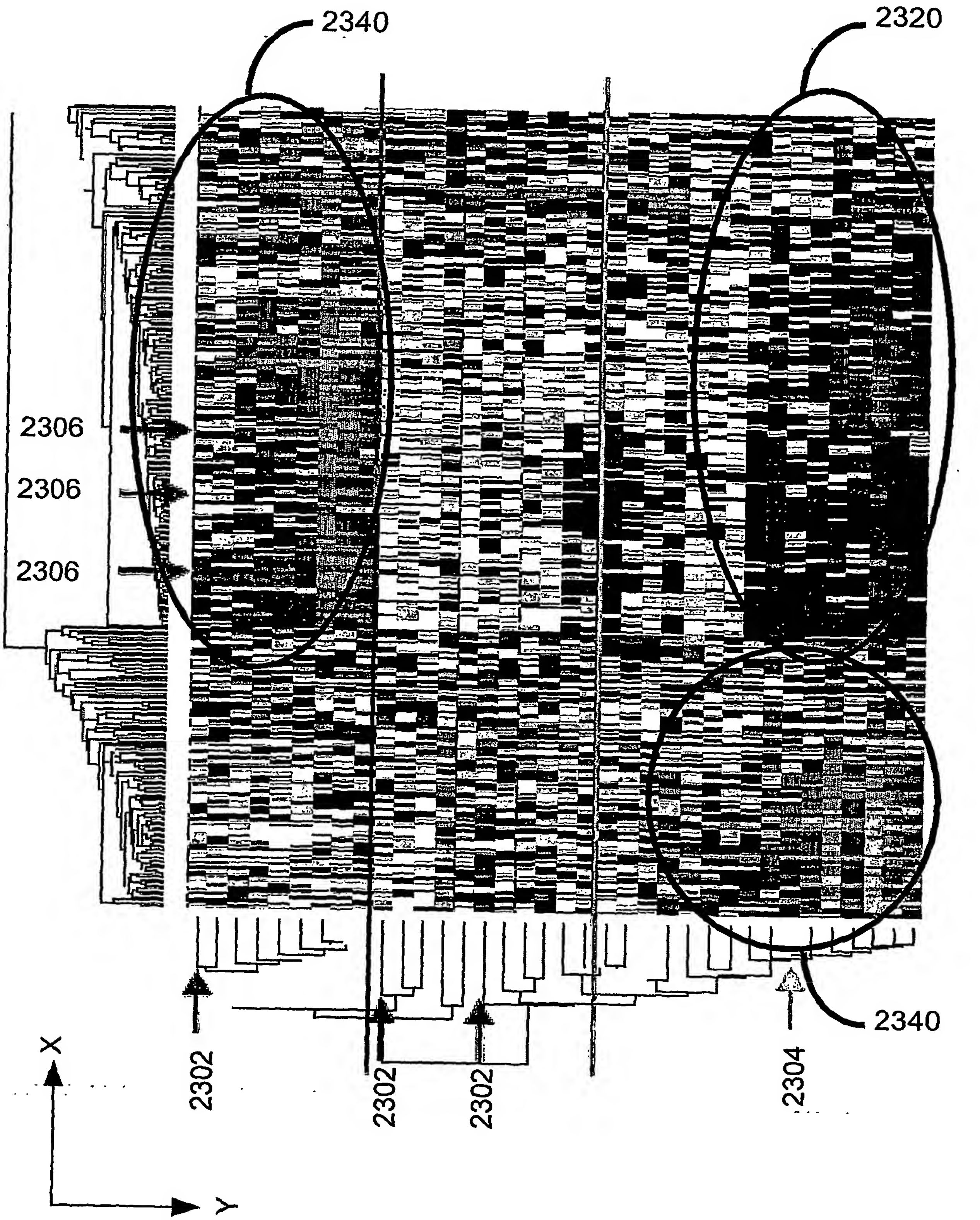



FIG. 23

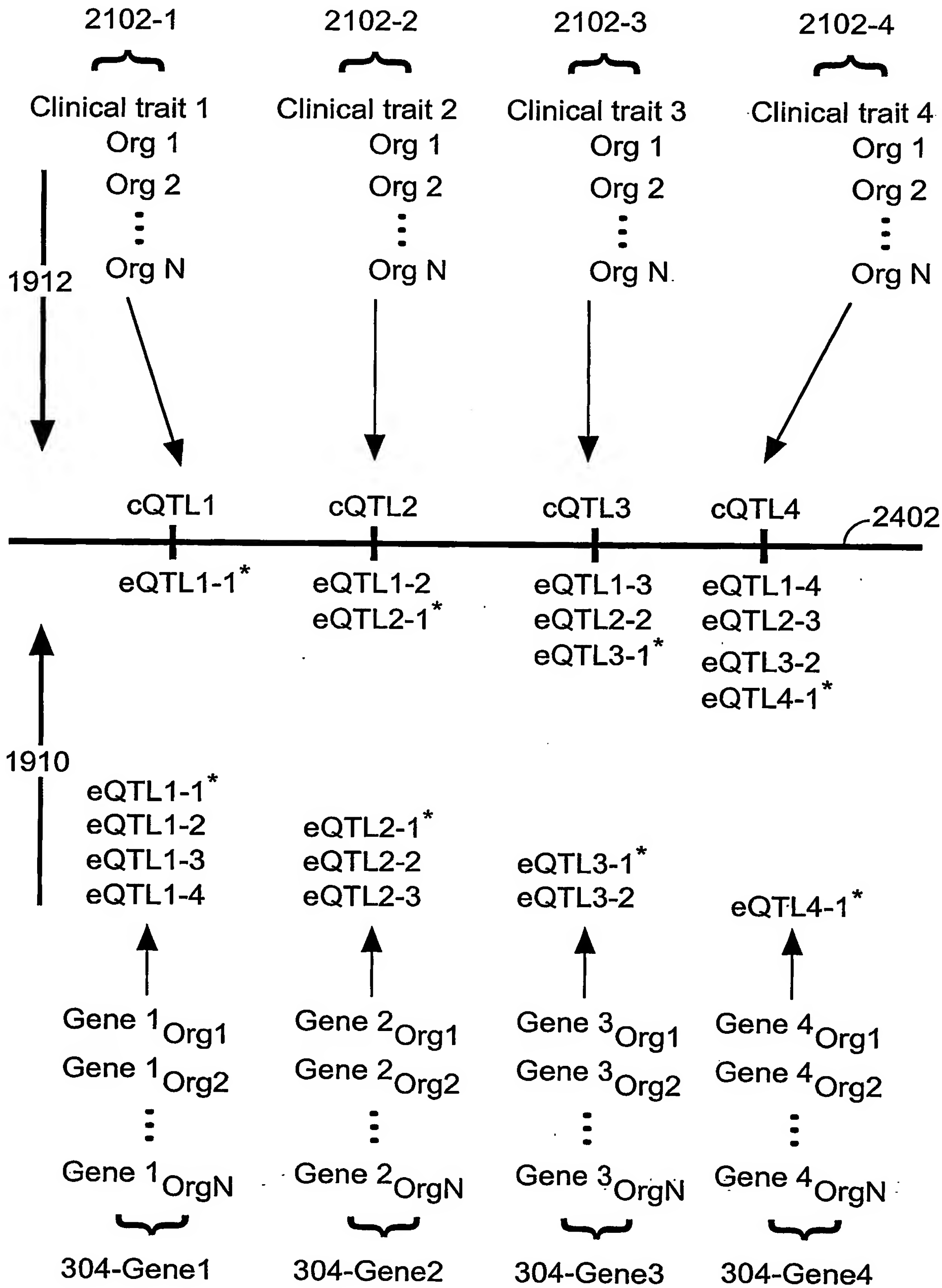

2102-1
2102-2
2102-3
2102-4
Clinical trait 1
Org 1
Org 2
Org N
Clinical trait 2
Org 1
Org 2
Org N
Clinical trait 3
Org 1
Org 2
Org N
Clinical trait 4
Org 1
Org 2
Org N
1912
cQTL1
cQTL2
cQTL3
cQTL4
2402
eQTL1-1*
eQTL1-2
eQTL2-1*
eQTL1-3
eQTL2-2
eQTL3-1*
eQTL1-4
eQTL2-3
eQTL3-2
eQTL4-1*
1910
eQTL1-1*
eQTL1-2
eQTL1-3
eQTL1-4
eQTL2-1*
eQTL2-2
eQTL2-3
eQTL3-1*
eQTL3-2
eQTL4-1*
Gene 1Org1
Gene 1Org2
Gene 1OrgN
Gene 2Org1
Gene 2Org2
Gene 2OrgN
Gene 3Org1
Gene 3Org2
Gene 3OrgN
Gene 4Org1
Gene 4Org2
Gene 4OrgN
304-Gene1
304-Gene2
304-Gene3
304-Gene4

FIG. 24

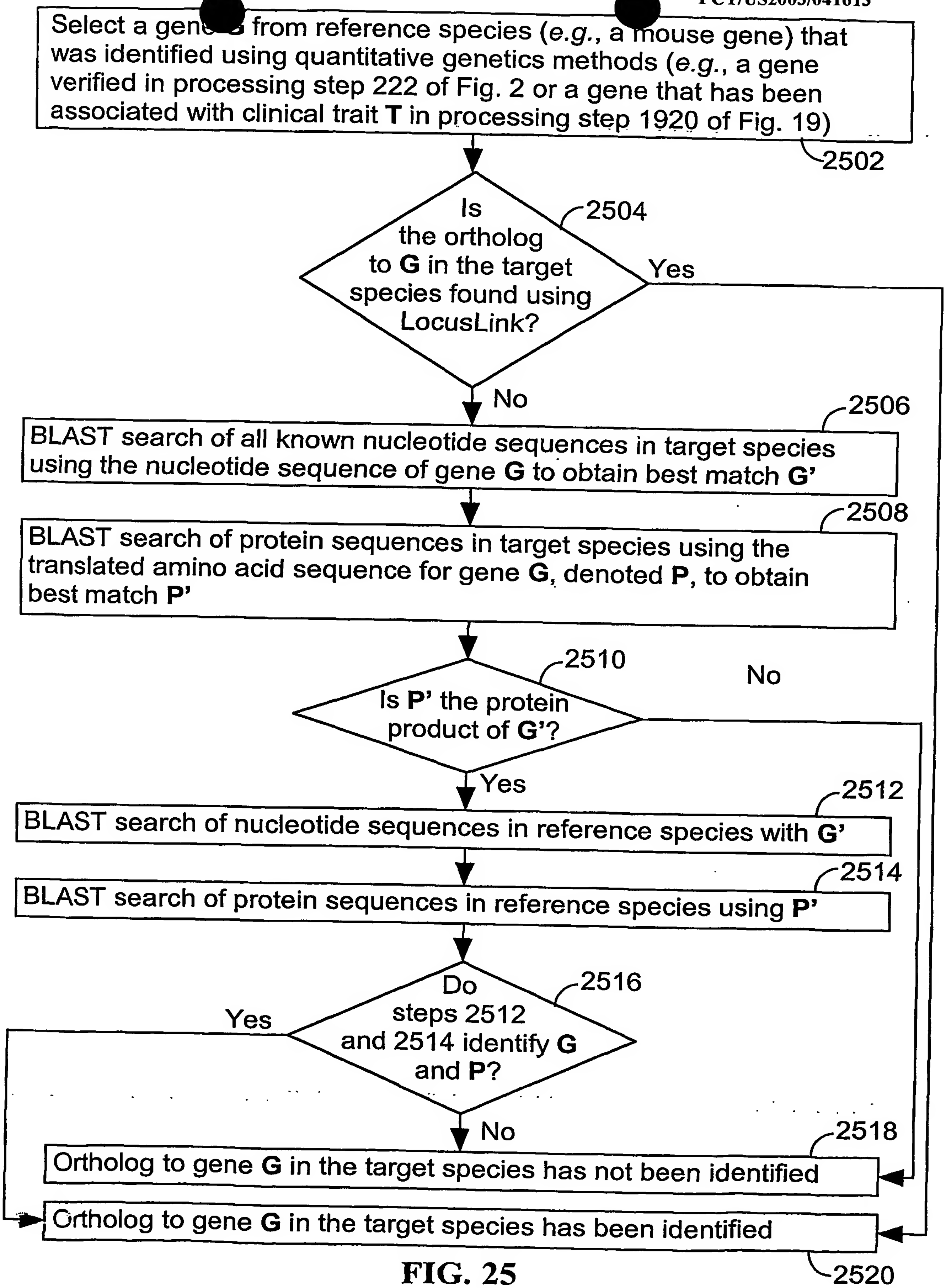

Select a gene G from reference species (e.g., a mouse gene) that was identified using quantitative genetics methods (e.g., a gene verified in processing step 222 of Fig. 2 or a gene that has been associated with clinical trait T in processing step 1920 of Fig. 19)
2502
Is the ortholog to G in the target species found using LocusLink?
2504
Yes
No
BLAST search of all known nucleotide sequences in target species using the nucleotide sequence of gene G to obtain best match G'
2506
BLAST search of protein sequences in target species using the translated amino acid sequence for gene G, denoted P, to obtain best match P'
2508
Is P' the protein product of G'?
2510
No
Yes
BLAST search of nucleotide sequences in reference species with G'
2512
BLAST search of protein sequences in reference species using P'
2514
Do steps 2512 and 2514 identify G and P?
2516
Yes
No
Ortholog to gene G in the target species has not been identified
2518
Ortholog to gene G in the target species has been identified
2520

FIG. 25

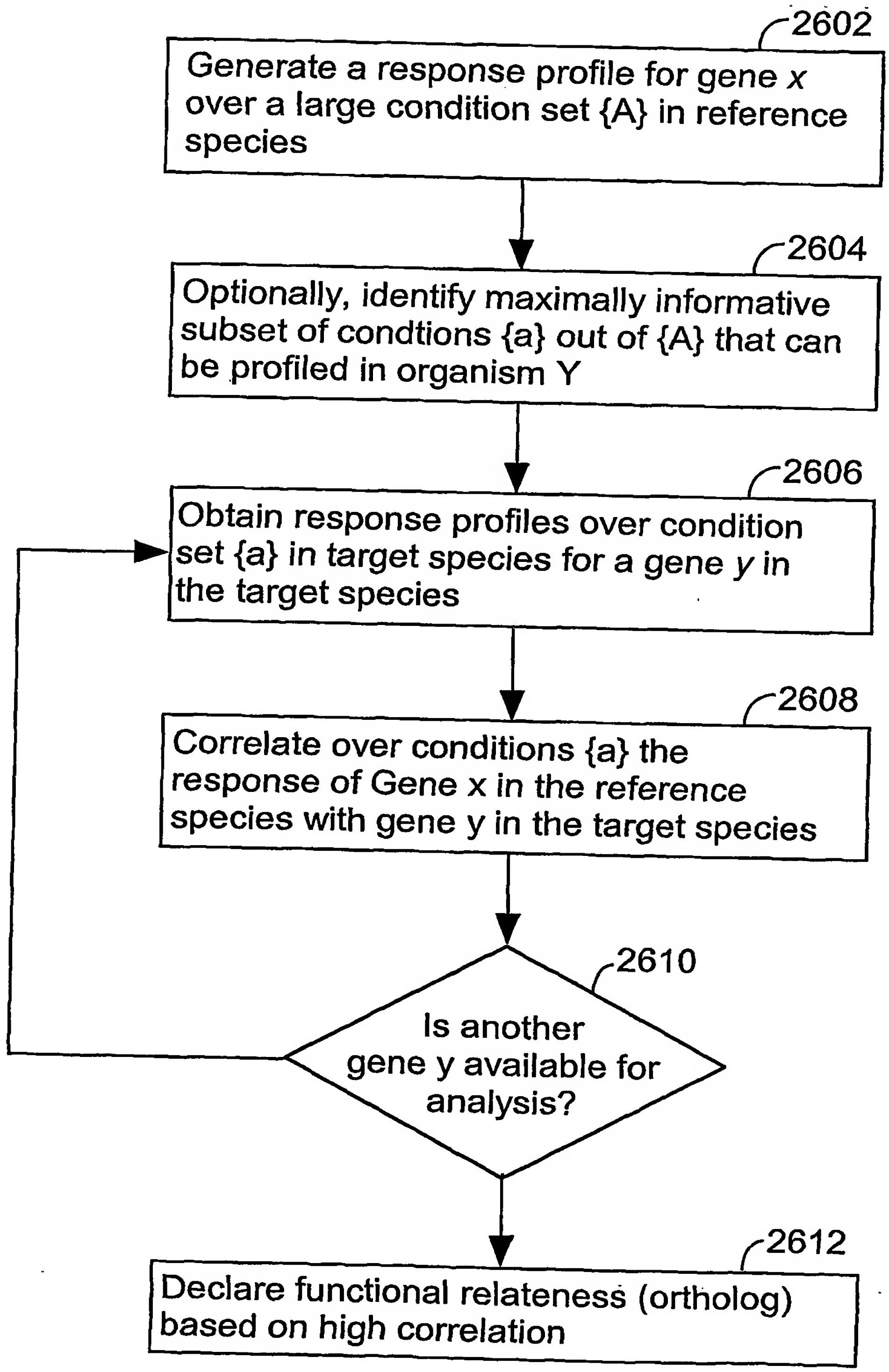

2602
Generate a response profile for gene x over a large condition set {A} in reference species
2604
Optionally, identify maximally informative subset of condtions {a} out of {A} that can be profiled in organism Y
2606
Obtain response profiles over condition set {a} in target species for a gene y in the target species
2608
Correlate over conditions {a} the response of Gene x in the reference species with gene y in the target species
2610
Is another gene y available for analysis?
2612
Declare functional relateness (ortholog) based on high correlation

FIG. 26

```
   1 GAGCTATTCG GCCTCTCTAG GCCGGCGGGT CCTCCGCTCC ATGGTCCTGT CTGTCAGCGC
  61 TGTGTCAGGA GGCCAGTGCC GAGGTCCGGT CGCGCTCCGA CGCTTCGACC CTCGAGCCGG
 121 TCGCGGGTAT CCCGGCGGCC GCGGGACGAT GGCGTGGTGG CACTGACAGG CGCGGGCGGC
 181 TGCCGAGCCC CGCGGCCGGC ATGGCGGGCC AGTTCCGCAG CTACGTGTGG GACCCGTTGC
 241 TAATCCTGTC GCAGATCGTA CTCATGCAGA CCGTCTACTA TGGCTCTCTG GGCCTGTGGC
 301 TGGCGCTGGT GGACGCGCTG GTGCGCAAGC CCGTCCCTGG ACCAGATGTT CGACGCGGAG
 361 ATCCTGGGCT TCTCCACCCC TCCAGGCCGG CTCTCAATGA TGTCCTTCGT CCTCAACGCC
 421 CTCACCTGTG CCCTGGGCTT GCTGTACTTC ATCCGGCGAG GGAAGCAGTG CCTGGATTTC
 481 ACTGTCACTG TGCATTTCTT TCACCTCCTG GGCTGCTGGC TCTACAGCTC CCGTTTCCCC
 541 TCGGCGCTGA CCTGGTGGCT GGTCCAGGCT GTGTGCATTG CACTCATGGC CGTCATCGGG
 601 GAGTACCTGT GCATGCGGAC GGAGCTCAAG GAGATCCCCC TCAGCTCAGC CCCTAAGTCC
 661 AATGTCTAGA GTTGGGCCCT TTGGACACTC TGCTGGCACT TGGGCCCCAT CACCTTGGGC
 721 TGCTCAGACC TCCAGATGGG GTCTGGCCCA AGTCTGAGCA GAACCCTGGA AATGTGAAGT
 781 CTGTTGGTGG AGAGATAATG AGGTCCCATC ATAAAGGCAG GTAGCAGCCA TGATCACAGA
 841 TGTAAGAATG GCCTCTGTCT GCCAAAGCCT TGATATCTGG AGGCCAGTAA GGGACCTCAT
 901 GGAGGGTAGT GGCAGATTTG GAACCATGTC ACATGAGCCA TCATACTGTC ACCAGCCTGT
 961 TATTTTAAAA AGAAAAAAAA AAAATCAAGG ATATCTGATT GGAATAAACC ACTCTTCTCG
1021 TTGTCTGTCT TATGCCCATG ACAGCCAGTA CCTTTGCTGT GTTGCCAAAC CACAGGGATT
1081 CTCTGTGGAG AAATACCTGA TTTCTGGGTC CATAGCCACA GAAAAAGATG TAGGTACAGA
1141 GTGCTAGGCT GCTGACAGGA CGTCGAGGGG AGGAGGCATC AAGCACAAGA AAAATGCATG
1201 GCCGTGCCGT TAGACACACA CACACACTTT TGTGTGTGTC CAGGACCCAT GACTGTCTCC
1261 CTCCAGTTCC CTGTATGGAC TCTGCCTTGC TGTTGTCACT CAGCACAGCC AGAGACAGGA
1321 CCCAGAGAAA ACCCCAGCAT CCCTCCCAGC CTTCCCTTCA TAATAAAAGC CATTGTCTGC
1381 TCTCTGGAAG TGAGCAGGCA GCCAGCTTCT ACTGGACCTC AACTGTGGCA GGAGTTTCTG
1441 TTTGCTGTCT TTTGAGTTCT GTGATAGGGA GGGTGTACTA AAGGTGCTGG AGGCTCACCC
1501 TGCTAAGCTT TCTTCCAAGT GGTTTCCTCA GGAAGGGCTG GCAGCTGTCC TTCCTAGGTA
1561 CATAAATACA CTATTTTCCA ATC
```

**Figure 27**

TCTAGGCCGGCAGCGCCTCTCCTCCATGGTCCTGTCTGTCAGCGCTGTTTTGGGAGCCCGCCGGTGAGGC
CGGGCCACGCTCAGACACTTCGATCGTCGAGTCTGTCACTGGGCATGGCGGGTCAGTTCCGCAGCTACGT
GTGGGACCCGCTGCTGATCCTGTCGCAGATCGTCCTCATGCAGACCGTGTATTACGGCTCGCTGGGCCTG
TGGCTGGCGCTGGTGGACGGGCTAGTGCGACAGCCCCTCGCTGGACCAGATGTTCGACGCCGAGATCCTG
GGCTTTTCCACCCCTCCAGGCCGGCTCTCCATGATGTCCTTCATCCTCAACGCCCTCACCTGTGCCCTGG
GCTTGCTGTACTTCATCCGGCGAGGAAAGCAGTGTCTGGATTTCACTGTCACTGTCCATTTCTTTCACCT
CCTGGGCTGCTGGTTCTACAGCTCCCGTTTCCCCTCGGCGCTGACCTGGTGGCTGGTCCAAGCCGTGTGC
ATTGCACTCATGGCTGTCATCGGGGAGTACCTGTGCATGCGGACGGAGCTCAAGGAGATACCCCTCAACT
CAGCCCCTAAATCCAATGTCTAGAATCAGGCCCTTTGGACATCCTGCTGACACTTGGGCCCCTTAACACC
TTGGGCTGCTCAGACCCTCCAGATGAGGTCCAGCCCAGATCTGAGAGGAACCCTGGAAATGTGAAGTCTC
TGTTGGTTTGGGAGAGATAGTGAGGGCCTGTCAAAGAAGGCAGGTAGCAGTCAGCATGACAGCTGCAAGA
ATGACCTCTGTCTGTTGAAGCCTTGGTATCTGAGAGGTCAGGAAGGGGACCTCTTTGAGGGTAATAACAG
AATTGGAACCATGCCACTCTTGAGCCACAATACCTGTCACCAGCCTGTTGTTTTAAGAGAGAAAAAAAT
CAAGGATATCTGATTGGAGCAAACCACTTCTTTAGTCATCTGTCTTACCCCCCTGGGACAGCTGTTACCT
TTGCAGTGTTGCCGAATCACAGCAGTTACCTTTGCAGTGTTGCCGAATCACAGCAGTTCTGTTGGAGAAA
CGCTTGGTTTCCGGATCCAGAGCCACAGAAAGAAATGTAGGTGTGAAGTATTAGGCTGCTGTCAGGGAGA
GGATGGCAGATGGAGGCATCAAGCACAAGGAAAATGCACAACCTGTGCCCTGTTATACACACGTTCATGT
GCACCCAAGAACCTATGACTTTCTTCCAGTTCCTTCTACCAGGTCCCCATCCTGCTGCCAGCTCTCAACA
TAGCAGGCCATAGGACCCAGAGAAGAATCCCAGCGTTGCTCAAAGTCTAACCATCATAAAGACACTGCCT
GTCTTCTAGGAATGACCAGGCACCCAGCTCCCACTGGACTCCAATTTTTTTCCTGCCTTATTTAGAATT
CTTTGGCGGGAAGGGTATGATGGGTTCCCAGAGACAAGAAGCCCAACCTTCTGGCCTGGGCTGTGCTGAT
AGTGCTGAGGGAGATAGGAATTTGCTGCTAAGATTTTTCTTTGGGGTGGAGTTTCCTCTGTGAGGGGCTT
GCAGCTATCCTTCCTGTGTATACAAATACAGTATTTTCCATGGTTCTGCCTGCACTTACTTTGTAATGCC
ACGGTTGAGATTGAGAGAGATCAGCGCAGCCAGGCAAGGGAACTTTAAAGAATTATTAGGCCACCTTCTC
CCTTTCCTGGACCCCAGAGTCATTCCTCCATTTGGTTAAAATACTCAGTGCAGGGAACTCTTACATCCTG
TCTCCTTCACTTGCAGCGTCCCCTGCTATGCCTCAGGTGAACCACATAATTCTTGGGTTTCCGTTCCTAC
TTGCTAGTGATTTCTGAACATGTTCAATGGAGCGGCACACAGTCTAGACCCACTTCCGCATTGAAACCTT
CACTGTTCCTCTTTGGTTTCTTCAGAGCTTTCCCAAGAGAGCTGTCAGTTTTCAGCTGTCAGTAACACAA
ATGAGTTTATGGTAACACAAATGAGTTTTGCTATCTCTCTGAGAAGCTCATCTGACCTCCTGACTCTCAG
CCCTACAGAGTAGGGAGTTGATGCTGACAGGATGAAGATTTAGGAATAAATATGCCTGGGAAGAGACTGG
GAAGGTTCTAGGGTGAGGCACCTCAGTAACTCATGGTACCTTGGCCAAGTTGGAAGGAAGCAGTTTGTTA
ATGAGGCACAGTAATCCTGGCTGCAGGGTCTAGGAGGTAAGACCAGCTGGGATGACCTTCCCTGGGTTAA
TCAATTTCCCTCTAGACAACACAAACTGCAGGCATGTGACTAACTTTGAAAGAACACCCATCATGTGGCT
GCTGTCACCCTTGACCAGCCGTGGTGGTGGTTACTCCATCTGTGGTTGGAGCGCCTCTTTGGGATTCACT
TCAAGGTCTTGTGCCTATTTTTCTGCATATCTTCTGTGATGACAAATCTCTGTCCCCTGAGTGTTAATTT
GATTTTTAGAAATGGCCAAAAGTCACGTGATCCAAACTTTTTTTCAGTAATATGGAGACTGAGCTGCATG
GTAGTTGGGGATCAAAAATATGTGACCTTAATGAGATTTTTATGATTTCTAAAGTAACAATAAAAGCAGT
TTTTAGAGTTGAGTTCCAGAGAGGGCAGGGCAATGGCAGTGACATGTTTGTCATTTTAATAATAAATAAC
ATCTATTGAGTGCTTAA

**Figure 28**

ATGGCGGGTCAGTTCCGCAGCTACGTGTGGGACCCGCTGCTGATCCTGTCGCAGATCGTCCTCATGCAGA
CCGTGTATTACGGCTCGCTGGGCCTGTGGCTGGCGCTGGTGGACGGGCTAGTGCGACAGCCCCTCGCTGG
ACCAGATGTTCGACGCCGAGATCCTGGGCTTTTCCACCCCTCCAGGCCGGCTCTCCATGATGTCCTTCAT
CCTCAACGCCCTCACCTGTGCCCTGGGCTTGCTGTACTTCATCCGGCGAGGAAAGCAGTGTCTGGATTTC
ACTGTCACTGTCCATTTCTTTCACCTCCTGGGCTGCTGGTTCTACAGCTCCCGTTTCCCCTCGGCGCTGA
CCTGGTGGCTGGTCCAAGCCGTGTGCATTGCACTCATGGCTGTCATCGGGGAGTACCTGTGCATGCGGAC
GGAGCTCAAGGAGATACCCCTCAACTCAGCCCC

**Figure 29**

MAGQFRSYVW DPLLILSQIV LMQTVYYGSL GLWLALVDAL VRSSPSLDQM FDAEILGFST
PPGRLSMMSF VLNALTCALG LLYFIRRGKQ CLDFTVTVHF FHLLGCWLYS SRFPSALTWW
LVQAVCIALM AVIGEYLCMR TELKEIPLSS APKSNV

**Figure 30A**

MALWACGWRW WTRWCAQPVP GPDVRRGDPG LLHPSRPALN DVLRPQRPHL CPGLAVLHPA
REAVPGFHCH CAFLSPPGLL ALQLPFPLGA DLVAGPGCVH CTHGRHRGVP VHADGAQGDP
PQLSP

**Figure 30B**

MAGQFRSYVW DPLLILSQIV LMQTVYYGSL GLWLALVDAL VRKPVPGPDV RRGDPGLLHP
SRPALNDVLR PQRPHLCPGL AVLHPAREAV PGFHCHCAFL SPPGLLALQL PFPLGADLVA
GPGCVHCTHG RHRGVPVHAD GAQGDPPQLS P

**Figure 30C**

MAGQFRSYVW DPLLILSQIV LMQTVYYGSL GLWLALVDAL VRSSPSLDQM FDAEILGFST
PPGRLSMMSF VLNALTCALG LLYFIRRGKQ CLDFTVTVHF FHLLGCWLYS SRFPSALTWW
LVQAVCIALM AVIGEYLCMR TELKEVPLSS APKSNV

**Figure 30D**

PFPGSRGPQL FGLSRPAGPP LHGPVCQRCV RRPVPRSGRA PTLRPSSRSR VSRRPRDDGV
VALTGAGGCR APRAGMAGQF RSYVWDPLLI LSQIVLMQTV YYGSLGLWWR WWTRWCAQPV
PGPDVRRGDP GLLHPSRPAL NDVLRPQRPH LCPGLAVLHP AREAVPGFHC HCAFLSPPGL
LALQLPFPLG ADLVAGPGCV HCTHGRHRGV PVHADGAQGD PPQLSP

**Figure 30E**

MAGQFRSYVW DPLLILSQIV LMQTVYYGSL GLWLALVDAL VRKPVPGPDV RRGDPGLLHP
SRPALNDVLR PQRPHLCPGL AVLHPAREAV PGFHCHCAFL SPPGLLALQL PFPLGADLVA
GPGCVHCTHG RHRGVPVHAD GAQGDPPQLS P

**Figure 30F**

MAGQFRSYVW DPLLILSQIV LMQTVYYGSL GLWLALVDGL VRQPLAGDPV RRRDPGLFHP
SRPALHDVLH PQRPHLCPGL AVLHPARKAV SGFHCHCPFL SPPGLLVLQL PFPLGADLVA
GPSRVHCTHG CHRGVPVHAD GAQGDTPQLS P

**Figure 31**

```
   1 GCACGAGGGC GGGCGCGCGC GTGGGCGCAG CGCGGAGCGG GGCCCATGGT GCGGCCGTGT
  61 CCGTCGGTCG GGCCGCGCGG GCGGCTCCGC GCGTGGCCCG GCGCTCGCGA CCTCGCCCCT
 121 GCGCTGCGGG CCCGGCCCGC CCGCTGCCGG CGCCTCCTCC CCCTGCCCCG GGGCGGCGCG
 181 GAGGCCGCGG GGAGCGCAGG GGGCGCGGCG GGCGGCGACA TGACGGACAG CATCCCGCTG
 241 CAGCCCGTGC GCCACAAGAA GCGGGTGGAC AGTAGGCCGC GCGCGGGGTG CTGTGAGTGG
 301 CTGAGATGTT GCGGTGGAGG GGAGCCCAGG CCCCGTACTG TCTGGTTGGG ACACCCCGAG
 361 AAGAGGGACC AGCGGTACCC TCGAAATGTC ATCAACAACC AGAAGTACAA TTTCTTCACA
 421 TTTCTTCCTG GGGTGTTGTT CAGCCAGTTC AGATACTTCT TCAACTTCTA CTTCCTGCTT
 481 CTCGCCTGCT CGCAGTTCGT CCCAGAGATG AGGCTTGGCG CCCTGTACAC CTACTGGGTT
 541 CCTCTGGGCT TCGTGCTGGC TGTCACCATC ATCCGTGAGG CAGTAGAGGA GATCCGATGT
 601 TATGTGCGTG ACAAGGAGAT GAACTCCCAG GTCTACAGCC GGCTCACGTC ACGAGGGACC
 661 GTGAAGGTGA AGAGTTCAAA CATCCAGGTG GGAGACCTCA TCCTTGTGGA AAAGAACCAG
 721 CGGGTCCCTG CTGACATGAT CTTCCTGAGG ACGTCAGAGA AAAACGGCTC TTGCTTCTTG
 781 CGCACGGATC AGCTGGATGG AGAGACAGAC TGGAAGCTTC GGCTCCCGGT GGCCTGCACA
 841 CAGAGGCTTC CCACGGCTGC TGACCTCCTG CAGATTCGGT CCTATGTGTA CGCTGAAAAA
 901 CCCAACATCG ACATTCACAA CTTCCTGGGG ACTTTCACCA GGGAAAACAG TGACCCTCCG
 961 ATCAGTGAGA GTCTGAGCAT TGAGAACACG CTGTGGGCCG GCACCGTCAT AGCATCAGGC
1021 ACTGTTGTAG GCGTTGTTCT CTACACTGGC AGAAAACTGC GGAGTGTCAT GAATACTTCC
1081 GACCCCAGAA GTAAGATTGG CCTGTTCGAC CTGGAGGTGA ACTGCCTCAC CAAAATCCTG
1141 TTTGGTGCGC TGGTGGTGGT GTCCCTGGTC ATGGTGGCCC TGCAGCACTT TGCCGGCCGC
1201 TGGTACCTGC AGATCATCCG CTTCCTGCTC CTGTTTTCCA ACATCATTCC TATCAGCTTG
1261 CGTGTGAACT TGGACATGGG CAAGATCGTG TACAGCTGGG TGATCCGCAG GGATTCCAAA
1321 ATCCCCGGGA CCGTGGTTCG TTCCAGCACA ATTCCTGAGC AGCTGGGCAG GATTTCGTAC
1381 TTGCTCACAG ACAAGACAGG AACCCTGACC CAGAATGAGA TGGTGTTCAA GCGGCTGCAC
1441 CTGGGTACGG TGGCCTACGG CCTGGACTCC ATGGACGAAG TGCAGAGTCA CATCTTCAGC
1501 ATTTACACCC AGCAATCCCA GGATCCACCT GCTCAGAAGG GCCCCACGGT CACCACCAAG
1561 GTCCGGAGGA CCATGAGCAG CCGTGTCCAC GAGGCTGTGA AGGCCATTGC ACTCTGCCAC
1621 AACGTGACAC CCGTGTACGA GTCCAATGGT GTGACGGACC AGGCTGAGGC TGAGAAGCAG
1681 TTTGAGGACT CCTGCCGAGT GTACCAGGCA TCCAGCCCGG ATGAGGTGGC TCTGGTCCAG
1741 TGGACAGAAA GTGTGGGACT GACGCTGGTG GGTCGAGACC AGTCCTCCAT GCAGCTGAGG
1801 ACCCCTGGTG ACCAGGTCCT GAATCTCACC ATCCTTCAGG TCTTCCCGTT CACCTATGAG
1861 AGCAAGCGGA TGGGCATCAT CGTGCGGGAT GAGTCCACGG GGGAAATCAC GTTCTACATG
1921 AAGGGAGCAG ACGTCGTCAT GGCTGGCATT GTCCAGTACA ACGACTGGCT GGAGGAGGAG
1981 TGTGGCAACA TGGCCCGGGA GGGACTACGT GTGCTGGTGG TAGCCAAGAA GTCCCTCACA
2041 GAGGAGCAGT ACCAACACTT TGAAGCCCGC TACGTCCAGG CTAAGCTGAG TGTGCATGAC
2101 CGCTCGCTGA AGGTGGCCAC GGTGATCGAG AGCTTGGAGA TGGAGATGGA GCTGCTGTGC
2161 CTGACTGGTG TGGAGGACCA GCTGCAGGCA GATGTCAGGC CCACGCTGGA GACGCTGCGC
2221 AACGCTGGCA TCAAGGTTTG GATGCTAACA GGGGACAAGC TGGAGACAGC CACGTGCACA
2281 GCCAAGAACG CACATCTGGT GACCAGAAAC CAAGATATCC ATGTTTTCCG ACTGGTGACC
2341 AACCGCGGGG AGGCCCACCT GGAGCTGAAT GCCTTCCGTA GGAAGCATGA CTGTGCCCTG
2401 GTCATCTCTG GAGACTCCCT GGAGGTTTGC CTCAAATACT ATGAGTACGA GTTCATGGAA
2461 CTGGCCTGCC AGTGCCCGGC TGTGGTGTGC TGCCGCTGTG CCCCAACCCA GAAGGCCCAG
2521 ATTGTTCGGC TGCTCCAAGA ACGCACCGGG AAACTCACCT GTGCAGTATG GGACGGAGGC
2581 AATGACGTCA GCATGATCCA GGAATCCGAC TGCGGCGTGG GCGTGGAGGG CAAGGAAGGG
2641 AAGCAGGCCT CGCTGGCAGC GGACTTCTCC ATCACCCAGT TCAAGCATCT CGGCCGCTTG
2701 CTCATGGTGC ACGGTCGGAA CAGCTACAAG CGCTCGGCGG CCCTCAGTCA GTTTGTGATC
2761 CACAGGAGCC TCTGCATCAG CACCATGCAG GCTGTCTTCT CGTCTGTGTT CTACTTTGCA
```

Figure 32A

2821 TCCGTTCCTC TCTACCAAGG CTTCCTGATC ATTGGGTATT CTACCATCTA CACGATGTTT
2881 CCCGTGTTCT CCCTGGTTTT GGACAAAGAC GTGAAGTCGG AAGTCGCCAT GTTGTATCCT
2941 GAGCTCTACA AGGACCTGCT TAAGGGGCGG CCACTGTCCT ACAAGACGTT CTTAATTTGG
3001 GTGTTAATCA GCATCTATCA AGGGAGCACC ATCATGTACG GGGCGCTGCT GCTGTTCGAG
3061 TCGGAGTTTG TACACATCGT GGCAATCTCC TTCACATCCC TCATCCTCAC TGAGCTACTG
3121 ATGGTGGCGC TCACCATCCA GACGTGGCAC TGGCTCATGA CAGTGGCCGA GCTACTCAGC
3181 CTGGCCTGCT ACATTGCCTC CCTGGTGTTC CTCCATGAGT TCATCGATGT CTACTTCATT
3241 GCCACCCTGT CATTCCTCTG GAAGGTGTCC GTCATCACCT TGGTCAGCTG TCTCCCCCTC
3301 TATGTCCTCA AGTACCTGCG GAGACGGTTC TCCCCACCCA GCTACTCGAA GCTCACTTCC
3361 TAAGCTGCAG GGCTGCCTCG GGCAGGGCCT CCGGCCTCCG GCGCTNTCCC CAGGAGGAGG
3421 TCAAGTTCCA CACGCACGAG CCGCCTCTGC TGGACGGTGC AGTCATGGCT GGCACATGAG
3481 GCTTCGCTGA GGCGACACTG GGCACCTAAT GGGGATGGAA CATTGGTGGA ACCGGAGGGA
3541 GGGACCTGAG AGCTGTACCT ATCAGAACCT TGGGTGCTAA GCTGTGCTGA GGGGGAAGAC
3601 GTGGGACCGG ATGGCCCGTC TGAGGTTTGT GGGGTCACTG TGCAAGCTTC CCTTATGGTT
3661 TGAACCTCTT GCCTGCAGCC CGGGG

**Figure 32B**

| LocusID | Org | Symbol | Description |
|---|---|---|---|
| ☐ 11981 | *Mm* | **Atp9a** | ATPase, class II, type 9A |
| More | *Mm* | | ATPase, class 2; ATPase 9A, p type; Class II; ATPase 9A, class II |

## Symbol and Name (MGI)

**Atp9a: ATPase, class II, type 9A**

**LocusID:** 11981

## Overview ?

| | |
|---|---|
| **Locus Type:** | gene with protein product, function known or inferred |
| **Product:** | ATPase, class 2 |
| **Alternate Symbols:** | IIa |
| **Alias:** | Class II<br>ATPase, class 2<br>ATPase 9A, p type<br>ATPase 9A, class II |

## Function Submit GeneRIF ?

Gene Ontology™:

| Term | Evidence | Source | Pub |
|---|---|---|---|
| • membrane | IEA | MGD | |
| • hydrolase | IEA | MGD | |
| • metabolism | IEA | MGD | |
| • ATP binding | IEA | MGD | |
| • cation transport | IEA | MGD | |
| • magnesium binding | IEA | MGD | |
| • integral membrane protein | IEA | MGD | |
| • plasma membrane cation-transporting ATPase | IEA | MGD | |

## Relationships ?

**Human Homology Maps:**

| | | | |
|---|---|---|---|
| NCBI vs. MGD | 20q13.11-13.2 | ATP9A | **Hs** |
| UCSC vs. MGD | 20q13.11-13.2 | ATP9A | **Hs** |

Figure 33

## NCBI Reference Sequences (RefSeq) ?

Category: NCBI Genome Annotation

| | | |
|---|---|---|
| Genomic Contig: | NT_011362 | gb sv mv ev mm |

Annotated transcripts/proteins for this locus:

| | | |
|---|---|---|
| Evidence: | supported by alignment with both mRNA and ESTs (27) | |
| Model mRNA: | XM_030577 | |
| Model Protein: | XP_030577 | BL |

## GenBank Sequences ?

| Nucleotide | Type | Protein | |
|---|---|---|---|
| AB014511 | m | BAA31586 | BL |
| AK025559 | m | | |
| AK026513 | m | | |
| BC016044 | m | AAH16044 | BL |

Figure 34

```
   1 MTDSIPLQPV RHKKRVDSRP RAGCCEWLRC CGGGEPRPRT VWLGHPEKRD QRYPRNVINN
  61 QKYNFFTFLP GVLFSQFRYF FNFYFLLLAC SQFVPEMRLG ALYTYWVPLG FVLAVTIIRE
 121 AVEEIRCYVR DKEMNSQVYS RLTSRGTVKV KSSNIQVGDL ILVEKNQRVP ADMIFLRTSE
 181 KNGSCFLRTD QLDGETDWKL RLPVACTQRL PTAADLLQIR SYVYAEKPNI DIHNFLGTFT
 241 RENSDPPISE SLSIENTLWA GTVIASGTVV GVVLYTGRKL RSVMNTSDPR SKIGLFDLEV
 301 NCLTKILFGA LVVVSLVMVA LQHFAGRWYL QIIRFLLLFS NIIPISLRVN LDMGKIVYSW
 361 VIRRDSKIPG TVVRSSTIPE QLGRISYLLT DKTGTLTQNE MVFKRLHLGT VAYGLDSMDE
 421 VQSHIFSIYT QQSQDPPAQK GPTVTTKVRR TMSSRVHEAV KAIALCHNVT PVYESNGVTD
 481 QAEAEKQFED SCRVYQASSP DEVALVQWTE SVGLTLVGRD QSSMQLRTPG DQVLNLTILQ
 541 VFPFTYESKR MGIIVRDEST GEITFYMKGA DVVMAGIVQY NDWLEEECGN MAREGLRVLV
 601 VAKKSLTEEQ YQHFEARYVQ AKLSVHDRSL KVATVIESLE MEMELLCLTG VEDQLQADVR
 661 PTLETLRNAG IKVWMLTGDK LETATCTAKN AHLVTRNQDI HVFRLVTNRG EAHLELNAFR
 721 RKHDCALVIS GDSLEVCLKY YEYEFMELAC QCPAVVCCRC APTQKAQIVR LLQERTGKLT
 781 CAVWDGGNDV SMIQESDCGV GVEGKEGKQA SLAADFSITQ FKHLGRLLMV HGRNSYKRSA
 841 ALSQFVIHRS LCISTMQAVF SSVFYFASVP LYQGFLIIGY STIYTMFPVF SLVLDKDVKS
 901 EVAMLYPELY KDLLKGRPLS YKTFLIWVLI SIYQGSTIMY GALLLFESEF VHIVAISFTS
 961 LILTELLMVA LTIQTWHWLM TVAELLSLAC YIASLVFLHE FIDVYFIATL SFLWKVSVIT
1021 LVSCLPLYVL KYLRRRFSPP SYSKLTS
```

**Figure 35**

```
   1 MTDNIPLQPV RQKKRMDSRP RAGCCEWLRC CGGGEARPRT VWLGHPEKRD QRYPRNVINN
  61 QKYNFFTFLP GVLFNQFKYF FNLYFLLLAC SQFVPEMRLG ALYTYWVPLG FVLAVTVIRE
 121 AVEEIRCYVR DKEVNSQVYS RLTARGTVKV KSSNIQVGDL IIVEKNQRVP ADMIFLRTSE
 181 KNGSCFLRTD QLDGETDWKL RLPVACTQRL PTAADLLQIR SYVYAEEPNI DIHNFVGTFT
 241 REDSDPPISE SLSIENTLWA GTVVASGTVV GVVLYTGREL RSVMNTSNPR SKIGLFDLEV
 301 NCLTKILFGA LVVVSLVMVA LQHFAGRWYL QIIRFLLLFS NIIPISLRVN LDMGKIVYSW
 361 VIRRDSKIPG TVVRSSTIPE QLGRISYLLT DKTGTLTQNE MIFKRLHLGT VAYGLDSMDE
 421 VQSHIFSIYT QQSQDPPAQK GPTLTTKVRR TMSSRVHEAV KAIALCHNVT PVYESNGVTD
 481 QAEAEKQYED SCRVYQASSP DEVALVQWTE SVGLTLVGRD QSSMQLRTPG DQILNFTILQ
 541 IFPFTYESKR MGIIVRDEST GEITFYMKGA DVVMAGIVQY NDWLEEECGN MAREGLRVLV
 601 VAKKSLAEEQ YQDFEARYVQ AKLSVHDRSL KVATVIESLE MEMELLCLTG VEDQLQADVR
 661 PTLETLRNAG IKVWMLTGDK LETATCTAKN AHLVTRNQDI HVFRLVTNRG EAHLELNAFR
 721 RKHDCALVIS GDSLEVCLKY YEYEFMELAC QCPAVVCCRC APTQKAQIVR LLQERTGKLT
 781 CAVGDGGNDV SMIQESDCGV GVEGKEGKQA SLAADFSITQ FKHLGRLLMV HGRNSYKRSA
 841 ALSQFVIHRS LCISTMQAVF SSVFYFASVP LYQGFLIIGY STIYTMFPVF SLVLDKDVKS
 901 EVAMLYPELY KDLLKGRPLS YKTFLIWVLI SIYQGSTIMY GALLLFESEF VHIVAISFTS
 961 LILTELLMVA LTIQTWHWLM TVAELLSLAC YIASLVFLHE FIDVYFIATL SFLWKVSVIT
1021 LVSCLPLYVL KYLRRRFSPP SYSKLTS
```

**Figure 36**

ATGACGGACAACATCCCGCTGCAGCCGGTGCGCCAGAAGAAGCGGATGGACAGCAGGCCC
CGCGCCGGGTGCTGCGAGTGGCTGAGATGCTGCGGTGGAGGGGAGGCCAGGCCCCGCACT
GTCTGGCTGGGGCACCCCGAGAAGAGAGACCAGAGGTATCCTCGGAATGTCATCAACAAT
CAGAAGTACAATTTCTTCACCTTTCTTCCTGGGGTGCTGTTCAACCAGTTCAAATACTTT
TTCAACCTCTATTTCTTACTTCTTGCCTGCTCTCAGTTTGTTCCCGAAATGAGACTTGGT
GCACTCTATACCTACTGGGTTCCCCTGGGCTTCGTGCTGGCCGTCACTGTCATCCGTGAG
GCGGTGGAGGAGATCCGATGCTACGTGCGGGACAAGGAAGTCAACTCCCAGGTCTACAGC
CGGCTCACAGCACGAGGCACAGTGAAGGTGAAGAGTTCTAACATCCAAGTTGGAGACCTT
ATCATCGTTGAAAAGAACCAGCGGGTCCCTGCCGACATGATCTTCCTGAGGACATCAGAA
AAAAACGGGTCATGCTTCTTGCGGACGGATCAGCTGGATGGGGAGACGGACTGGAAGCTG
CGGCTTCCCGTGGCCTGCACGCAGAGGCTCCCCACGGCCGCCGACCTTCTTCAGATTCGA
TCGTATGTGTACGCAGAAGAGCCAAATATTGACATTCACAACTTCGTGGGAACTTTTACC
CGAGAAGACAGCGACCCCCCGATCAGCGAGAGCCTGAGCATAGAGAACACGCTGTGGGCT
GGCACTGTGGTCGCATCAGGTACTGTTGTGGGTGTTGTTCTTTACACTGGCAGAGAACTC
CGGAGTGTCATGAATACCTCAAATCCCCGAAGTAAGATCGGCCTGTTCGACTTGGAAGTG
AACTGCCTCACCAAGATCCTCTTTGGTGCCCTGGTGGTGGTCTCGCTGGTCATGGTTGCC
CTTCAGCACTTTGCAGGCCGTTGGTACCTGCAGATCATCCGCTTCCTCCTCTTGTTTTCC
AACATCATCCCCATTAGTTTGCGTGTGAACCTGGACATGGGCAAGATCGTGTACAGCTGG
GTGATTCGAAGGGACTCGAAAATCCCCGGGACCGTGGTTCGCTCCAGCACGATTCCTGAG
CAGCTGGGCAGGATTTCGTACTTACTCACAGACAAGACAGGCACTCTTACCCAGAACGAG
ATGATTTTCAAACGGCTCCATCTCGGAACAGTAGCCTACGGCCTCGACTCAATGGACGAA
GTACAAAGCCACATTTTCAGCATTTACACCCAGCAATCCCAGGACCCACCGGCTCAGAAG
GGCCCAACGCTCACCACTAAGGTCCGGCGGACCATGAGCAGCCGCGTGCACGAAGCCGTG
AAGGCCATCGCGCTCTGCCACAACGTGACTCCCGTGTATGAGTCCAACGGTGTGACTGAT
CAGGCTGAGGCCGAGAAGCAGTACGAAGACTCCTGCCGCGTATACCAGGCATCCAGCCCC
GATGAGGTGGCCCTGGTACAGTGGACGGAAAGTGTGGGCTTAACCCTGGTGGGCCGAGAC
CAGTCTTCCATGCAGCTGAGGACCCCTGGCGACCAGATCCTGAACTTCACCATCCTACAG
ATCTTCCCTTTCACCTATGAAAGCAAACGTATGGGCATCATCGTGCGGGATGAATCAACT
GGAGAAATTACGTTTTACATGAAGGGAGCAGATGTGGTCATGGCTGGCATTGTGCAGTAC
AATGACTGGTTGGAGGAAGAGTGTGGCAACATGGCCCGAGAAGGGCTGCGGGTGCTCGTG
GTGGCAAAGAAGTCTCTTGCAGAGGAGCAGTATCAGGACTTTGAAGCCCGCTACGTCCAG
GCCAAGCTGAGTGTGCACGACCGCTCCCTCAAAGTGGCCACGGTGATCGAGAGCCTGGAG
ATGGAGATGGAACTGCTGTGCCTGACGGGCGTGGAGGACCAGCTGCAGGCAGATGTGCGG
CCCACGCTGGAGACCCTGAGGAATGCTGGCATCAAGGTTTGGATGCTGACAGGGGACAAG
CTGGAGACAGCTACGTGCACAGCGAAGAATGCACATCTGGTGACCAGAAACCAAGACATC
CACGTTTTTCGGCTGGTGACCAACCGCGGGGAGGCTCACCTCGAGCTGAACGCCTTCCGC
AGGAAGCATGATTGTGCCCTGGTCATCTCGGGAGACTCCCTGGAGGTTTGCCTCAAGTAC
TATGAGTACGAGTTCATGGAGCTGGCCTGCCAGTGCCCGGCCGTAGTCTGCTGCCGATGT
GCCCCCACCCAGAAGGCCCAGATCGTGCGCCTGCTTCAGGAGCGCACGGGCAAGCTCACC
TGTGCAGTAGGGGACGGAGGCAATGACGTCAGCATGATTCAGGAATCTGACTGCGGCGTG
GGAGTGGAAGGAAAGGAAGGAAAACAGGCTTCGTTGGCTGCAGACTTCTCCATCACTCAA
TTTAAGCATCTTGGCCGGTTGCTTATGGTGCATGGCCGGAACAGCTACAAGCGGTCAGCC
GCCCTCAGCCAGTTCGTGATTCACAGGAGCCTCTGTATCAGCACCATGCAGGCTGTCTTT
TCCTCCGTGTTTTACTTTGCCTCCGTCCCTCTCTATCAAGGATTCCTCATCATTGGGTAC
TCCACAATTTACACCATGTTTCCTGTGTTTTCTCTGGTCCTGGACAAAGATGTCAAATCG
GAAGTTGCCATGCTGTATCCTGAGCTCTACAAGGATCTTCTCAAGGGACGGCCGTTGTCC
TACAAGACATTCTTAATATGGGTTTTGATTAGCATCTATCAAGGGAGCACCATCATGTAC
GGGGCGCTGCTGCTGTTTGAGTCGGAGTTCGTGCACATCGTGGCCATCTCCTTCACCTCG
CTGATCCTCACCGAGCTGCTCATGGTGGCGCTGACCATCCAGACCTGGCACTGGCTCATG
ACAGTGGCGGAGCTGCTCAGCCTGGCCTGCTACATCGCCTCCCTGGTGTTCTTACACGAG
TTCATCGATGTGTACTTCATCGCCACCTTGTCATTCTTGTGGAAAGTCTCCGTCATCACT
CTGGTCAGCTGCCTCCCCCTCTATGTCCTCAAGTACCTGCGAAGACGGTTCTCTCCCCCC
AGCTACTCAAAGCTCACATCA

Figure 37

```
   1 GGGAAGCTGT TGCGCACCAC TTAGCTGGGA AGTGCGTTGC TCCCTGTTTC CCAGCCCACC
  61 CGAGATGGCC CCCAAAGTCT CGGATTCCGT GGAACAGCTC CGCGCTGCCG GCAACCAGAA
 121 CTTCCGCAAT GGCCAGTACG GCGAAGCTCG GCGCTGTACG AGCGCGCACT GCGGCTGCTG
 181 CAGGCGCGAG GTAGGAACCC GCCCCACGTT TCCCTCCGGG CCTGCGTCCT CCACCCGCAT
 241 CCCCGCACCG GGCCTCCCGT TGGCCCAGCC TCCCTGGTTT TCCCTTCCCC GCGTCCAGCC
 301 GCCGCACCAG GCCCTCCCAG GGCTTGACCC CGCGATTCTT TCCGTCCCTG GCCGCCTAGC
 361 CGCGGCCCGG TCTACCATCA CCACCCCCCA CCACCCCCAG GCCAGTCGGC TGCGGGCCTT
 421 AAGGGCACGC ATCCGCTGCT TCCACCCGGA AGCTGTTGCG CACTCCTCGG CGGGGAACGG
 481 AGGTGGTCCT TGTTTGCCGG CCTCCCGGGA TGGCCCCCAA ACTCTCAGAC TCTGTGGAAG
 541 AGCTCCGCGC AGCCGGCAAC CAGAGTTTCC GCAACGGACA GTACGCCGAG GCTTCGGCGC
 601 TGTACGAGCG CGCGCTGCGA CTGCTGCAGG CGCGAGGTTC TGCAGACCCC GAAGAAGAAA
 661 GTGTTCTGTA CTCCAACCGT GCAGCGTGCT ACTTGAAGGA TGGGAACTGC ACAGATTGCA
 721 TCAAAGATTG CACTTCCGCG CTGGCCTTGG TTCCCTTCAG CATCAAGCCC TTGCTGCGCA
 781 GAGCATCTGC ATATGAAGCC CTGGAGAAGT ACGCCCTGGC CTACGTTGAC TATAAGACTG
 841 TGCTGCAGAT CGATAACAGT GTGGCATCCG CCCTGGAAGG CATCAACAGA ATAACCAGAG
 901 CTCTCATGGA CTCCCTGGGA CCTGAGTGGC GCCTGAAGCT GCCCCCTATC CCTGTGGTGC
 961 CTGTTTCAGC CCAGAAGAGA TGGAATTCCT TGCCTTCAGA TAACCACAAA GAGACAGCTA
1021 AAACCAAATC CAAAGAAGCC ACAGCTACGA AGAGCAGAGT GCCTTCTGCT GGGGATGTGG
1081 AGAGAGCCAA AGCTCTGAAG GAAGAAGGCA ATGACCTTGT AAAGAAGGGC AACCATAAGA
1141 AAGCTATTGA GAAGTACAGT GAGAGCCTCT TGTGTAGTAG CCTGGAGTCT GCCACATACA
1201 GCAACAGAGC GCTCTGTCAC CTGGTCCTGA AGCAGTACAA GGAGGCAGTA AAGGACTGCA
1261 CAGAAGCCCT CAAGCTGGAT GGGAAGAATG TAAAGGCGTT TTACAGACGG GCTCAAGCCT
1321 ACAAGGCACT CAAGGACTAT AAGTCAAGCC TTTCGGATAT CAGCAGCCTC CTACAAATTG
1381 AACCCAGGAA TGGCCCTGCA CAGAAGTTAC GGCAGGAAGT TAACCAGAAC ATGAACTAAA
1441 CCGTAGAGGG CAACAGGGAC CCTGAACTTG ACCTTCCCAG AGAAGCCAGG GCCTCCCTTG
1501 CATCTGCCCC AATGCCCAGC ATGCCGCCAA GGGAGTGCAA AATCAACCCC ACTTTGACTC
1561 CTTGGAGAGG TAGCAGCCTT TCACCTGACA CATTTTACTT GTTCAGATTA AGTCCATTAC
1621 AGACAAGCAC AGGACTCTTT TTTTTTTTCT TCTTTTTTTT TTCCAGAAAG GTCCCCACTA
1681 GAGGTTTTTG TTTTGTTTTA TTTTTAATTT AAAAAAGCGT GACGCCAACA GCCCTGGCCT
1741 CATTCGCTTG CTTCTGCCTG GCCCTTGTCA ACACAGTCCT TGGCAACTGT CCCTGACCCA
1801 GATATGCACA GACTGGGTGC CTGTGACTTC CTCTGCCGCC ATAGCTCTGC AGTTCACCTG
1861 AGTGCTGACA GGCTAGAAGT GCTTGCTCGT CCGCAGCCAC AGCGGCCTGT TGAGCTGGTT
1921 CTCCAAGGCT GCCTGCCATC TCCTCGAGGA GACAGCTGCT GTCTGCACCC TGTCCTTGAC
1981 ACAGTGTCCT GTGTTGAGCC CCAGTGCCTT TAGTCCAGGC CCTTTGTGGG AAGGCAGAGC
2041 CTAACCCTTG GAGGCTCTGT GTTGTTGCCT TCTGTCTGAG CTACCTACGA TGTTCAAAGA
2101 GCCCAGATTC CTCCTGCAAT GGGGAGAGAG GCCTCCTTGA GATTAGTGTC CCTCCAGTCT
2161 GAGCAGGAAC TTAACCTTTT CCCCCATAGC AGCAGCCCCT CGGGCTCCTT TGTTTTGTTT
2221 TGTTTTGTTA ATATGTTGGA GTTAATTGAA CTGATTTTAT TGAAGTGTGT GTTGCTGTTG
2281 CATTAAAAGG TTTTCTTCTA TG
```

**Figure 38**

Click to Display mRNA-Genomic Alignments (spanning 17338 bps)

PUB UNIGENE HOMOL MGI MGC

**Mus musculus Official Gene Symbol and Name (MGI)**

**2610100K07Rik: RIKEN cDNA 2610100K07 gene**

**LocusID:** 67145

**Overview** Submit GeneRIF ?

| | |
|---|---|
| **Locus Type:** | gene with protein product, function unknown |
| **Product:** | RIKEN cDNA 2610100K07 |

**Relationships** ?

**Human Homology Maps:**

| | | | |
|---|---|---|---|
| NCBI vs. MGD | 20q12-q13.1 | TOM34 | **Mm Hs** |
| UCSC vs. MGD | 20q12-q13.1 | TOM34 | **Mm Hs** |
| NCBI vs. MGD | 20q12-q13.1 | TOMM34 | **Mm Hs** |

**Map Information** ?

| | | |
|---|---|---|
| **Chromosome:** | 2 | **mv** |
| **Cytogenetic:** | 2 | RefSeq |

**NCBI Reference Sequences (RefSeq)** ?

**Category:** PREDICTED

| | | |
|---|---|---|
| **mRNA:** | NM_025996 | |
| **Protein:** | NP_080272 RIKEN cDNA 2610100K07 | **BL** |
| **Domains:** | TPR Domain | score: 84 |

| | | |
|---|---|---|
| **mRNA:** | NM_025996 | |
| **Protein:** | NP_080272 RIKEN cDNA 2610100K07 | |
| **Domains:** | TPR Domain | score: 84 |
| | Tetratricopeptide repeats | score: 88 |

Figure 39A

HOMOL | GDB | eI | UCSC

## *Homo sapiens* Official Gene Symbol and Name (HGNC)

**TOMM34: translocase of outer mitochondrial membrane 34**

**LocusID:** 10953

## Overview

**Proteome Summary:** Subunit of the translocase of the outer mitochondrial membrane; component of the mitochondrial protein import complex

| | |
|---|---|
| **Locus Type:** | gene with protein product, function known or inferred |
| **Product:** | translocase of outer mitochondrial membrane 34 |
| **Alternate Symbols:** | TOM34, HTOM34P |
| **Alias:** | outer mitochondrial membrane translocase (34kD) |

## Function Submit GeneRIF (All Pubs)

**GeneRIF: Gene References into Function:**

| | |
|---|---|
| 11913975 | • Tom34 unlike Tom20 does not interact with the leader sequences of mitochondrial precursor proteins |
| 11913976 | • Yeast two-hybrid screening identifies binding partners of human Tom34 that have ATPase activity and form a complex with Tom34 in the cytosol |

## NCBI Reference Sequences (RefSeq)

**Category: PROVISIONAL**

| | | | |
|---|---|---|---|
| **mRNA:** | NM_006809 | | |
| **Protein:** | NP_006800 translocase of outer mitochondrial membrane 34 | | BL |
| **Domains:** | TPR Domain | score: 86 | |
| | Tetratricopeptide repeats | score: 87 | |
| **GenBank Source:** | BC007423 | | |

**Category: NCBI Genome Annotation**

| | | |
|---|---|---|
| **Genomic Contig:** | NT_011362 | gb sv mv ev mm |

**Annotated transcripts/proteins for this locus:**

| | | | |
|---|---|---|---|
| **Evidence:** | supported by alignment with both mRNA and ESTs (37) | | |
| **Model mRNA:** | XM_029822 | | |
| **Model Protein:** | XP_029822 | | BL |
| **Domains:** | TPR Domain | score: 86 | |
| | Tetratricopeptide repeats | score: 89 | |

Figure 39B

MAPKLSDSVE ELRAAGNQSF RNGQYAEASA LYERALRLLQ ARGSADPEEE SVLYSNRAAC
YLKDGNCTDC IKDCTSALAL VPFSIKPLLR RASAYEALEK YALAYVDYKT VLQIDNSVAS
ALEGINRITR ALMDSLGPEW RLKLPPIPVV PVSAQKRWNS LPSDNHKETA KTKSKEATAT
KSRVPSAGDV ERAKALKEEG NDLVKKGNHK KAIEKYSESL LCSSLESATY SNRALCHLVL
KQYKEAVKDC TEALKLDGKN VKAFYRRAQA YKALKDYKSS LSDISSLLQI EPRNGPAQKL
RQEVNQNMN

**Figure 40**

MAPKFPDSVEELRAAGNESFRNGQYAEASALYGRALRVLQAQGSSDPEEESVLYSNRAACHLKDGNCRDC
IKDCTSALALVPFSIKPLLRRASAYEALEKYPMAYVDYKTVLQIDDNVTSAVEGINRMTRALMDSLGPEW
RLKLPSIPLVPVSAQKRWNSLPSENHKEMAKSKSKETTATKNRVPSAGDVEKARVLKEEGNELVKKGNHK
KAIEKYSESLLCSNLESATYSNRALCYLVLKQYTEAVKDCTEALKLDGKNVKAFYRRAQAHKALKDYKSS
FADISNLLQIEPRNGPAQKLRQEVKQNLH

**Figure 41**

1 GGCACGAGGC ACCACACGGG GGAGGAAGGA AGGAGCTCCC AACTCGCCGG CCTGGCCACG
61 GGATGGCCCC CAAATTCCCA GACTCTGTGG AGGAGCTCCG CGCCGCCGGC AATGAGAGTT
121 TCCGCAACGG CCAGTACGCC GAGGCCTCCG CGCTCTACGG CCGCGCGCTG CGGGTGCTGC
181 AGGCGCAAGG TTCTTCAGAC CCAGAAGAAG AAAGTGTTCT CTACTCCAAC CGAGCAGCAT
241 GTCACTTGAA GGATGGAAAC TGCAGAGACT GCATCAAAGA TTGCACTTCA GCACTGGCCT
301 TGGTTCCCTT CAGCATTAAG CCCCTGCTGC GGCGAGCATC TGCTTATGAG GCTCTGGAGA
361 AGTACCCTAT GGCCTATGTT GACTATAAGA CTGTGCTGCA GATTGATGAT AATGTGACGT
421 CAGCCGTAGA AGGCATCAAC AGAATGACCA GAGCTCTCAT GGACTCGCTT GGGCCTGAGT
481 GGCGCCTGAA GCTGCCCTCA ATCCCCTTGG TGCCTGTTTC AGCTCAGAAG AGGTGGAATT
541 CCTTGCCTTC GGAGAACCAC AAAGAGATGG CTAAAAGCAA ATCCAAAGAA ACCACAGCTA
601 CAAAGAACAG AGTGCCTTCT GCTGGGGATG TGGAGAAAGC CAGAGTTCTG AAGGAAGAAG
661 GCAATGAGCT TGTAAAGAAG GGAAACCATA AGAAAGCTAT TGAGAAGTAC AGTGAAAGCC
721 TCTTGTGTAG TAACCTGGAA TCTGCCACGT ACAGCAACAG AGCACTCTGC TATTTGGTCC
781 TGAAGCAGTA CACAGAAGCA GTGAAGGACT GCACAGAAGC CCTCAAGCTG GATGGAAAGA
841 ACGTGAAGGC ATTCTACAGA CGGGCTCAAG CCCACAAAGC ACTCAAGGAC TATAAATCCA
901 GCTTTGCAGA CATCAGCAAC CTCCTACAGA TTGAGCCTAG GAATGGTCCT GCACAGAAGT
961 TGCGGCAGGA AGTGAAGCAG AACCTACACT AAAAACCCAA CAGGGCAACT GGAACCCCTG
1021 CCTGACCTTA CCCAGAGAAG CCATGGGCCA CCTGCTCTGT GCCCGCTCCT GAAACCCAGC
1081 ATGCCCCAAG TGAGCTCTGA AGCCCCCTCC TCAATCCCTT GATGGCCTCC CACCCTGTAA
1141 GAGGCTTTGC TTGTTCAAAT TAAACTCAGT GTAGTCAAAC ACAGACATGG TTGTTGCACC
1201 AGAAAGGTCC CCACTAGAGC TAAGCGTGAA GCTGAAGCTC TGTCCCTATT CCCCCAGCCC
1261 AGCTAGCTGA TCACACCAAC AGATCCTCAT CAGCAAAGCA TTTGGCTTTG TCCTGCCCAA
1321 GTGGGCTGCA GACTGAGTGC TGCCCTTGTA GCTTCCCCAG ACCCCAACTC ACTGCAGTTC
1381 ATCTGAACAA CCTGAGCTCC TGGGCCGGGG TGGAAGGAGG GGGATAAACC TAAGGCCCTG
1441 ATCCAAAGCA GCCTGTTGAG CTGGTTCTCC AGGGCTGCAG TCTCTCCAGG TGTACAGCTG
1501 CTGTCCCTGC CCTGTCCTGT CCTTGCACAG TCTCCTATGT CTGAGCCCCA GTGCCTTCTG
1561 TTCGGGCCCT CCTTTGGTGG GAAGGCAGAG CCCTGACCCT TGAATGGTTG TCCTTGACTC
1621 TGTGCTGCTG CCTTCTGCAG AGAGGCACCT AAGCTGTTTA AAGAGCCCAG TGATTGTGGC
1681 TGCTCCTCCT AGAGGTGGGA GGGGGCAAGA GGCCTCCTTG GTCAGTGTCC ATGCTTTCTG
1741 GGCAGGGACT TGGTTTTTTG TTCCAACAGT GGCCTTCTCC GGGCTTCATA GTTCTTTGTA
1801 ATATGTTGAA GTTAATTTGA ATTGACTGAT TTTGTTGAAC TGTGTGTTTA AGCTGTTGCA
1861 TTAAAAAGCT TTCTTCTACA TCAATATCTG CTGTGCTTTC ATTTATGCCT TTTCAGCTTT
1921 GCACCTGGAA CTCTGTAGTA ATAATAAAAG TTATTGCTTA TTGGGCATTC AAAAAAAAAA
1981 AAAAAAAA

**Figure 42**

MAPKVSDSVE QLRAAGNQNF RNGQYGEASA LYERALRLLQ ARGSADPEEE SVLYSNRAAC
YLKDGNCTDC IKDCTSALAL VPFSIKPLLR RASAYEALEK YALAYVDYKT VLQIDNSVAS
ALEGINRITR ALMDSLGPEW RLKLPPIPVV PVSAQKRWNS LPSDNHKETA KTKSKEATAT
KSRVPSAGDV ERAKALKEEG NDLVKKGNHK KAIEKYSESL LCSSLESATY SNRALCHLVL
KQYKEAVKDC TEALKLDGKN VKAFYRRAQA YKALKDYKSS LSDISSLLQI EPRNGPAQKL
RQEVNQNMN

**Figure 43**

MAPKFPDSVE ELRAAGNESF RNGQYAEASA LYGRALRVLQ AQGSSDPEEE SVLYSNRAAC
HWKNGNCRDC IKDCTSALAL VPFSIKPLLR RASAYEALEK YPMAYVDYKT VLQIDDNVTS
AVEGINRMTR ALMDSLGPEW RLKLPSFPLV PVSAQKRWNF LPSENHKEMA KSKSKETTAT
KNRVPSAGDV EKARVLKEEG NELVKKGNHK KAIEKYSESL LCSNLESATY SNRALCYLVL
KQYTEAVKDC TEALKLDGKN VKAFYRRAQA HKALKDYKSS FADISNLLQI EPRNGPAQKL
RQEVKQNLH

**Figure 44**

```
   1 GACTGGCTGG TGCGGGAAAT ATGCAGGAGA AAAGTCTTTG CATAATGTAG AGCGAGCCGT
  61 GGGGCTCCGG GAGCGGCGCC CCAAGGTCTG GGGCCATGAA CGCGAGCGTG GAAGGAGACA
 121 CCTTTTCTGG ATCGATGCAA ATCCCAGGAG GCACCACGGT CGTGGTGGAG CTGGCACCGG
 181 ACATCCACAT CTGCGGCCTC TGTAAGCAGC ACTTCAGCAA TCTGGATGCC TTTGTGGCCC
 241 ACAAACAGAG CGGCTGCCAG CTGACTACCA CGCCGGTGAC AGCCCCCAGC ACGGTCCAGT
 301 TTGTGGCAGA GGAGACAGAG CCTGCCACCC AGACCACCAC AACGACCATC AGTTCAGAGA
 361 CTCAGACTAT CACAGTTTCA GCTCCAGAGT TCGTCTTTGA ACATGGCTAC CAAACTTACC
 421 TGCCCACGGA GAGCACTGAC AACCAGACAG CCACCGTGAT CTCTCTCCCC ACCAAGTCAC
 481 GCACCAAAAA GCCCACAGCA CCCCCTGCTC AGAAGAGACT CGGCTGCTGC TATCCAGGTT
 541 GCCAGTTCAA GACCGCCTAT GGCATGAAGG ACATGGAGCG ACACCTGAAG ATCCACACCG
 601 GTGACAAACC CCACAAGTGT GAGGTGTGCG GGAAGTGCTT CAGCCGGAAG GACAAGCTGA
 661 AGACGCACAT GCGCTGCCAC ACGGGCGTCA AGCCCTACAA GTGCAAGACG TGCGACTACG
 721 CGGCGGCGGA CAGCAGCAGC CTTAACAAGC ACCTGCGCAT CCACTCGGAC GAGCGACCCT
 781 TCAAGTGCCA GATCTGTCCC TACGCCAGCC GCAACTCCAG CCAGCTCACC GTGCACCTGC
 841 GCTCGCACAC GGGGGACGCC CCCTTCCAGT GCTGGCTCTG TAGTGCCAAG TTCAAAATCA
 901 GCTCGGACTT GAAAAGGCAC ATGCGTGTGC ACTCGGGGGA GAAGCCTTTC AAGTGCGAAT
 961 TCTGCAATGT CCGCTGTACC ATGAAGGGGA ACCTCAAATC GCACATCCGC ATCAAGCACA
1021 GTGGGAATAA CTTCAAGTGT CCGCACTGCG ACTTCCTGGG TGACAGCAAA TCCACCCTGC
1081 GGAAGCACAG TCGCCTGCAC CAGTCGGAGC ACCCGGAGAA GTGTCCCGAG TGCAGCTACT
1141 CCTGTTCCAG CAAGGCCGCG CTGCGCGTGC ACGAGCGCAT CCACTGCACC GAGCGCCCGT
1201 TCAAGTGCAG CTACTGCAGC TTCGATACCA AGCAACCCAG CAACCTGAGC AAGCACATGA
1261 AGAAGTTCCA CGCCGACATG CTCAAGAACG AGGCTCCGGA GAAGAAGGAG AGCGGCAGGC
1321 AGAGCAGCCG GCAGGTGGCC AGGCTGGATG CCAAGAAGAC GTTCCACTGC GACATCTGTG
1381 ACGCCTCGTT TATGCGGGAG GACTCGCTCC GCAGCCACAA ACGGCAGCAC AGTGAGTACC
1441 ACAGTAAGAA CTCGGACGTG ACTGTAGTAC AGCTTCACCT TGAACCCAGC AAGCAGCCGC
1501 TGCGCCCCTC ACCGTAGAGC AAATCCAGGT CCCCCTCCAG TCCAGCCAGG TGCCCCAGTT
1561 CAGCGAGGGG AGGGTCAAGA TCATCGTGGG GCATTACAGG TGCCTCAGAC GAACCGCCAT
1621 AGTCCAAGCG GCCGCAGCTG CCGTCAACAT TGTGCCCCCC ACCCTGGTAG CCCAGACCCC
1681 AGAGGAGATC CCAGGGAACG GCCGGCTACA GATCCTTCGC CAGGTCAGTC TCATTGCCCC
1741 TCCTCAGTCC TCCGGGTGTC CCGGCGAAGC AGGTGCCCTG AGTCAGCCAA CTGTCCTGCT
1801 GACCACCCAT GATCAGACGG CAGGGGCCGC CCTGCAGCAG GCTCTGATCC CCACCACCCC
1861 GGTTGGGACC CAGGAAGGCA CGGGAAACCA GACATTCATT GCCAGTTCGG GCATCGTGCT
1921 CGGACTTGGA AGGCCTTAAG CTCTATTCAG GAGGGAACGA CGGAAGTGAC TGTGGTGAGC
1981 GATGGGGACC AGAGCATCGC AGTGGCCACC ACGGCACCCT CTATCTTCTC TACCCAGCAG
2041 GAACTGCCCA AGCAGACTTA CTCCATCATC CACGGGGCGG CACACCCCGC CCTGCTCTGT
2101 CCCGCCGACT CCATTCCTGA TTAGTCTGGA GGGAGGGGTG ACAGACAAGA CAAACTGCGA
2161 GAGGAGTACT GTGAGAGGCT CCTGGTCCCG CATAAATAAT TGTATTTTAT ACAGTTTATG
2221 TAATTTTTTA ACAGGGTATC AAGCTGGAGA CCATTCTCCC TCAAGCTCTT GTTGATTGTG
2281 TCTTAATGGT TACCAAGGCT GATTCCAATG TGGAGTTGGA ATTCACCACA GTAGGACTGA
2341 ATACATTCGT TTGTTTTTCC ATGTTTAGGA TTTAATTTTT TTCAACTGGA ATAAAGGAGT
2401 TTGGGATTTG GGTTAAAAAA
```

**Figure 45**

1 GAGTCCTCCC CGCCTCGCAG AGTTGGGAGA AGGCAGGGTG GGGGGTGTGG AAAAATAAAA
61 GGAAAAGTCC TTGCACCATG TAGATCAGCG TCCCCCACTT TGGCATCCCG GCCGGCCGGG
121 GACCTCCCAG TCTGCGGCCA TGAACGCGAG CAGCGAGGGC GAGAGCTTCG CGGGCTCGGT
181 GCAAATTCCA GGTGGCACAA CGGTGCTGGT GGAGCTGACT CCCGACATCC ATATCTGCGG
241 CATCTGCAAG CAGCAGTTTA ACAACCTGGA TGCCTTTGTA GCTCACAAGC AAAGTGGCTG
301 CCAGCTGACA GGCACATCCG CAGCAGCCCC CAGCACGGTC CAGTTTGTAT CGGAGGAAAC
361 AGTGCCTGCC ACCCAGACTC AGACCACCAC CAGAACCATC ACCTCGGAGA CCCAGACAAT
421 CACAGTTTCA GCTCCAGAAT TTGTTTTTGA ACATGGCTAT CAAACTTACC TGCCCACGGA
481 AAGTAATGAA AACCAGACAG CCACTGTCAT CTCTCTCCCT GCCAAGTCAC GCACCAAAAA
541 GCCCACAACA CCACCTGCTC AGAAAAGGCT TAACTGTTGC TATCCAGGTT GCACCAAAAA
601 GACTGCTTAT GGCATGAAGG ACATGGAGCG GCATTTAAAA ATTCACACGG GCCAATTCAA
661 CCATAAGTGT GAAGTCTGTG GCAAGTGCTT TAGCCGGAAA GACAAGCTGA AAACTCACAT
721 GCGGTGCCAC ACGGGCGTGA AGCCCTACAA GTGTAAGACG TGTGACTACG CCGCTGCCGA
781 CAGCAGCAGC CTCAACAAGC ACCTGAGGAT CCACTCGGAC GAGCGGCCCT TCAAATGCCA
841 GATCTGCCCC TACGCCAGCC GCAACTCCAG CCAGCTCACT GTCCACCTGC GATCCCACAC
901 GGGGGACGCC CCCTTCCAGT GCTGGCTCTG TAGCGCCAAG TTCAAAATCA GCTCGGACTT
961 GAAAAGGCAC ATGCGGGTGC ACTCGGGGGA GAAGCCTTTC AAGTGCGAGT TCTGCAATGT
1021 CCGCTGCACC ATGAAGGGGA ACCTCAAGTC GCACATCCGT ATCAAGCACA GCGGGAATAA
1081 CTTCAAGTGT CCTCATTGCG ACTTCCTGGG TGACAGCAAA GCCACCCTCC GGAAGCACAG
1141 CCGCGTGCAC CAGTCGGAGC ATCCTGAGAA GTGCTCGGAA TGCAGCTACT CCTGCTCCAG
1201 CAAGGCCGCC CTGCGCATCC ACGAGCGTAT CCACTGCACC GACCGCCCTT TCAAGTGCAA
1261 CTACTGCAGC TTCGACACCA AACAGCCCAG CAACCTGAGC AAGCACATGA AGAAGTTCCA
1321 CGGGGACATG GTTAAGACTG AGGCTCTAGA GAGGAAGGAC ACCGGCAGGC AGAGCAGCCG
1381 GCAGGTGGCC AAGCTGGATG CCAAGAAGAG TTTCCACTGC GATATATGCG ATGCCTCCTT
1441 CATGCGGGAG GACTCGCTCC GCAGCCACAA GAGACAGCAC AGTGAGTACA ATGAGAGTAA
1501 GAACTCGGAC GTGACCGTTC TCCAGTTTCA GATCGACCCC AGCAAGCAGC CCGCCACGCC
1561 CCTCACTGTG GGACACCTCC AGGTGCCCCT CCAGCCCAGC CAAGTGCCCC AGTTCAGCGA
1621 GGGAAGAGTC AAAATCATCG TTGGGCATCA GGTGCCCCAG GCGAACACCA TCGTCCAGGC
1681 TGCCGCTGCT GCAGTGAACA TCGTCCCGCC TGCCTTGGTG GCCCAGAACC CAGAGGAACT
1741 CCCAGGGAAC AGCCGGCTGC AGATCCTGCG CCAGGTCAGT CTGATCGCCC CCCCTCAGTC
1801 CTCGCGGTGT CCGAGCGAGG CGGGCGCAAT GACCCAGCCG GCTGTCCTGC TGACCACCCA
1861 CGAGCAGACG GACGGAGCCA CTCTGCACCA GACTCTCATC CCCACGGCCT CAGGTGGCCC
1921 CCAGGAAGGC TCTGGCAATC AAACTTTCAT TACCAGTTCG GGTATTACTT GCACTGACTT
1981 TGAAGGCCTA AACGCCTTGA TTCAGGAGGG GACAGCAGAA GTGACAGTGG TGAGCGATGG
2041 AGGCCAGAAC ATCGCAGTGG CCACCACAGC GCCACCGGTC TTCTCCTCCT CTTCCCAGCA
2101 AGAACTACCC AAGCAGACCT ACTCCATCAT TCAAGGGGCA GCCCATCCAG CTTTGCTCTG
2161 TCCCGCCGAC TCCATTCCAG ATTAGTGCTT AAAAAACAAA AGGAGTGGGG GAAAGGAATT
2221 GAGAAAAAGA AATCTTAAGT AGAATTCTCT AAAAGGTTTG CTCTTAATGT TTTCTTTGTT
2281 TTGTTTTGTT TTTGAGACGG AGTCTCGCTC TGTTTCCCAG GCTGGAGTGC AGTGGCGCTA
2341 TCTTGGCTCA CTGCAACGTC CGCCTCCCAG GTTCAAGCGA TTCTCATGCC TCGGCCCTCC
2401 GAGTAGCTGG GACCACAGGT GTACGACATC ATGACTGGCT AATTTTTGTA TATTTAATAG
2461 AGGCGGGGTT TCATCATGTT GAACTCCTGA CCTCAAGTGA TCTGCCCACC TCAGCCTCCC
2521 AAAGTGCTGG GATTACAGGT GTGAGCCACC ATGCCTGGCC GTGGTTTGCT CTTAATGTTT
2581 TTAAGGATGG TTGTGAATCC CCCTGGCCCC ATAATAAATT GTAATTTTAT ACTGCTTACT
2641 ATAATTTTTT TAACACTGTA ACAACTTTGA GACCACCTCT GAATCGTCGC ATTATAACTG
2701 TTGTAGAATC TTAAATGTTA CCAAGATGAT TCCAATGAGG GGTTGGAATT AAATGCATTA
2761 AGTAGTGAAC TCATGTGTTT GTTTCCAACT TGATTTTCCA ACTCTAATAA AGGTTTCTGT
2821 CCATCTTATT ACATTTGTGT AGTAAATGGT ACTTCCCAGC CTCTCTTTTG CCCCATTCTG
2881 GAATACTCCC CAGAGTTTGG GGGTGTTCAT GTTTTATACA TGTAAGTCTG TTGGCATGAA
2941 GGACCATTTT CTACATAATA TGACATGGAT ACTTGACCCA AAAAAAATGT TTAGTGCTAA
3001 TGAGCAGAAA ATGAATGGTT CCATAATAAA TTGATATCTG ATTAAAAT

**Figure 46**

MNASVEGDTF SGSMQIPGGT TVVVELAPDI HICGLCKQHF SNLDAFVAHK QSGCQLTTTP
VTAPSTVQFV AEETEPATQT TTTTISSETQ TITVSAPEFV FEHGYQTYLP TESTDNQTAT
VISLPTKSRT KKPTAPPAQK RLGCCYPGCQ FKTAYGMKDM ERHLKIHTGD KPHKCEVCGK
CFSRKDKLKT HMRCHTGVKP YKCKTCDYAA ADSSSLNKHL RIHSDERPFK CQICPYASRN
SSQLTVHLRS HTGDAPFQCW LCSAKFKISS DLKRHMRVHS GEKPFKCEFC NVRCTMKGNL
KSHIRIKHSG NNFKCPHCDF LGDSKSTLRK HSRLHQSEHP EKCPECSYSC SSKAALRVHE
RIHCTERPFK CSYCSFDTKQ PSNLSKHMKK FHADMLKNEA PEKKESGRQS SRQVARLDAK
KTFHCDICDA SFMREDSLRS HKRQHSEYHS KNSDVTVVQL HLEPSKQPLR PSP

**Figure 47**

MNASVEGDTF SGSMQIPGGT TVVVELAPDI HICGLCKQHF SNLDAFVAHK QSGCQLTTTP
VTAPSTVQFV AEETEPATQT TTTTISSETQ TITVSAPEFV FEHGYQTYLP TESTDNQTAT
VISLPTKSRT KKPTAPPAQK RLGCCYPGCQ FKTAYGMKDM ERHLKIHTGD KPHKCEVCGK
CFSRKDKLKT HMRCHTGVKP YKCKTCDYAA ADSSSLNKHL RIHSDERPFK CQICPYASRN
SSQLTVHLRS HTGDAPFQCW LCSAKFKISS DLKRHMRVHS GEKPFKCEFC NVRCTMKGNL
KSHIRIKHSG NNFKCPHCDF LGDSKSTLRK HSRLHQSEHP EKCPECSYSC SSKAALRVHE
RIHCTERPFK CSYCSFDTKQ PSNLSKHMKK FHADMLKNEA PEKKESGRQS SRQVARLDAK
KTFHCDICDA SFMREDSLRS HKRQHSEYHS KNSDVTVVQL HLEPSKQPLR PSP

**Figure 48**

MNASVEGDTF SGSMQIPGGT TVLVELAPDI HICGLCKQHF SNLDAFVAHK QSGCQLTTTP
VTAPSTVQFV AEETEPATQT TTTTISSETQ TITVSAPEFV FEHGYQTYLP TESTDNQTAT
VISLPTKSRT KKPTAPPAQK RLGCCYPGCQ FKTAYGMKDM ERHLKIHTGD KPHKCEVCGK
CFSRKDKLKT HMRCHTGVKP YKCKTCDYAA ADSSSLNKHL RIHSDERPFK CQICPYASRN
SSQLTVHLRS HTASVLENDV QKPAGLPAEE SDAQQAPAVT LSLEAKERTA TLGERTFNCR
YPGCHFKTVH GMKDLDRHLR IHTGDKPHKC EFCDKCFSRK DNLTMHMRCH TSVKPHKCHL
CDYAAVDSSS LKKHLRIHSD ERPYKCQLCP YASRNSSQLT VHLRSHTGDT PFQCWLCSAK
FKISSDLKRH MIVHSGEKPF KCEFCDVRCT MKANLKSHIR IKHTFKCLHC AFQGRDRADL
LEHSRLHQAD HPEKCPECSY SCSNPAALRV HSRVHCTDRP FKCDFCSFDT KRPSSLAKHI
DKVHREGAKT ENRAPPGKDG PGESGPHHVP NVSTQRAFGC DKCGASFVRD DSLRCHRKQH
SDWGENKNSN LVTFPSEGIA TGQLGPLVSV GQLESTLEPS HDL

**Figure 49**

MNASVEGDTF SGSMQIPGGT TVLVELAPDI HICGLCKQHF SNLDAFVAHK QSGCQLTTTP
VTAPSTVQFV AEETEPATQT TTTTISSETQ TITVSAPEFV FEHGYQTYLP TESTDNQTAT
VISLPTKSRT KKPTAPPAQK RLGCCYPGCQ FKTAYGMKDM ERHLKIHTGD KPHKCEVCGK
CFSRKDKLKT HMRCHTGVKP YKCKTCDYAA ADSSSLNKHL RIHSDERPFK CQICPYASRN
SSQLTVHLRS HTAWRCDCLG STKPWVPSLV TT

**Figure 50**

MNASSEGESF AGSVQIPGGT TVLVELTPDI HICGICKQQF NNLDAFVAHK QSGCQLTGTS
AAAPSTVQFV SEETVPATQT QTTTRTITSE TQTITVSAPE FVFEHGYQTY LPTESNENQT
ATVISLPAKS RTKKPTTPPA QKRLNCCYPG CQFKTAYGMK DMERHLKIHT GDKPHKCEVC
GKCFSRKDKL KTHMRCHTGV KPYKCKTCDY AAADSSSLNK HLRIHSDERP FKCQICPYAS
RNSSQLTVHL RSHTGDAPFQ CWLCSAKFKI SSDLKRHMRV HSGEKPFKCE FCNVRCTMKG
NLKSHIRIKH SGNNFKCPHC DFLGDSKATL RKHSRVHQSE HPEKCSECSY SCSSKAALRI
HERIHCTDRP FKCNYCSFDT KQPSNLSKHM KKFHGDMVKT EALERKDTGR QSSRQVAKLD
AKKSFHCDIC DASFMREDSL RSHKRQHSEY NESKNSDVTV LQFQIDPSKQ PATPLTVGHL
QVPLQPSQVP QFSEGRVKII VGHQVPQANT IVQAAAAAVN IVPPALVAQN PEELPGNSRL
QILRQVSLIA PPQSSRCPSE AGAMTQPAVL LTTHEQTDGA TLHQTLIPTA SGGPQEGSGN
QTFITSSGIT CTDFEGLNAL IQEGTAEVTV VSDGGQNIAV ATTAPPVFSS SSQQELPKQT
YSIIQGAAHP ALLCPADSIP D

**Figure 51**

MNASSEGESF AGSVQIPGGT TVLVELTPDI HICGICKQQF NNLDAFVAHK QSGCQLTGTS
AAAPSTVQFV SEETVPATQT QTTTRTITSE TQTITGCQFK TAYGMKDMER HLKIHTGDKP
HKCEVCGKCF SRKDKLKTHM RCHTGVKPYK CKTCDYAAAD SSSLNKHLRI HSDERPFKCQ
ICPYASRNSS QLTVHLRSHT GDAPFQCWLC SAKFKISSDL KRHMRVHSGE KPFKCEFCNV
RCTMKGNLKS HIRIKHSGNN FKCPHCDFLG DSKATLRKHS RVHQSEHPEK CSECSYSCSS
KAALRIHERI HCTDRPFKCN YCSFDTKQPS NLSKHMKKFH GDMVKTEALE RKDTGRQSSR
QVAKLDAKKS FHCDICDASF MREDSLRSHK RQHSEYSESK NSDVTVLQFQ IDPSKQPATP
LTVGHLQVPL QPSQVPQFSE GRVKIIVGHQ VPQANTIVQA AAAAVNIVPP ALVAQNPEEL
PGNSRLQILR QVSLIAPPQS SRCPSEAGAM TQPAVLLTTH EQTDGATLHQ TLIPTASGGP
QEGSGNQTFI TSSGITCTDF EGLNALIQEG TAEVTVVSDG GQNIAVATTA PPVFSSSSQQ
ELPKQTYSII QGAAHPALLC PADSIPD

**Figure 52**

MNASSEGESF AGSVQIPGGT TVLVELTPDI HICGICKQQF NNLDAFVAHK QSGCQLTGTS
AAAPSTVQFV SEETVPATQT QTTTRTITSE TQTITVSAPE FVFEHGYQTY LPTESNENQT
ATVISLPAKS RTKKPTTPPA QKRLNCCYPG CQFKTAYGMK DMERHLKIHT GDKPHKCEVC
GKCFSRKDKL KTHMRCHTGV KPYKCKTCDY AAADSSSLNK HLRIHSDERP FKCQICPYAS
RNSSQLTVHL RSHTGDAPFQ CWLCSAKFKI SSDLKRHMRV HSGEKPFKCE FCNVRCTMKG
NLKSHIRIKH SGNNFKCPHC DFLGDSKATL RKHSRVHQSE HPEKCSECSY SCSSKAALRI
HERIHCTDRP FKCNYCSFDT KQPSNLSKHM KKFHGDMVKT EALERKDTGR QSSRQVAKLD
AKKSFHCDIC DASFMREDSL RSHKRQHSEY SESKNSDVTV LQFQIDPSKQ PATPLTVGHL
QVPLQPSQVP QFSEGRVKII VGHQVPQANT IVQAAAAAVN IVPPALVAQN PEELPGNSRL
QILRQVSLIA PPQSSRCPSE AGAMTQPAVL LTTHEQTDGA TLHQTLIPTA SGGPQEGSGN
QTFITSSGIT CTDFEGLNAL IQEGTAEVTV VSDGGQNIAV ATTAPPVFSS SSQQELPKQT
YSIIQGAAHP ALLCPADSIP D

**Figure 53**

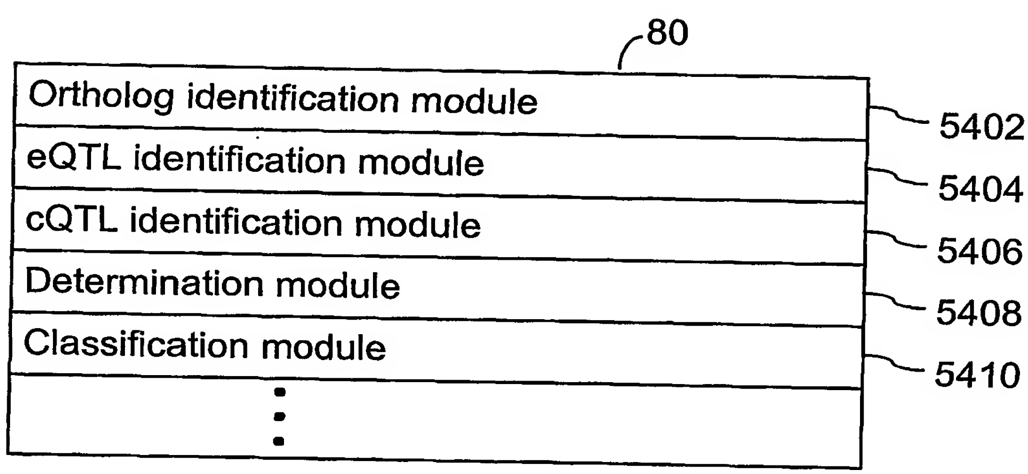
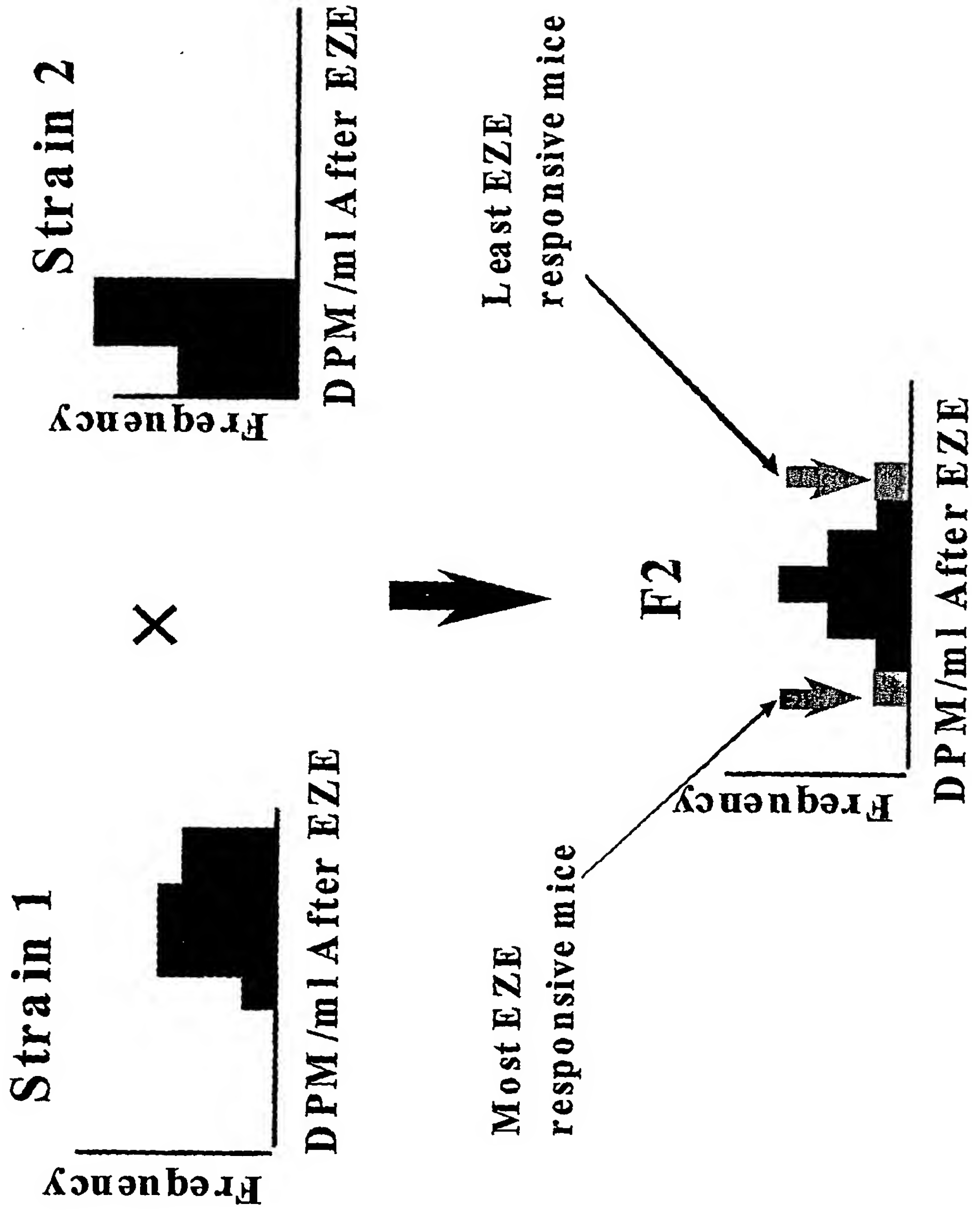

80
Ortholog identification module
eQTL identification module
cQTL identification module
Determination module
Classification module
5402
5404
5406
5408
5410

FIG. 54



Strain 1
Frequency
DPM/ml After EZE
×
Strain 2
Frequency
DPM/ml After EZE
F2
Frequency
DPM/ml After EZE
Most EZE
responsive mice
Least EZE
responsive mice


Fig. 55

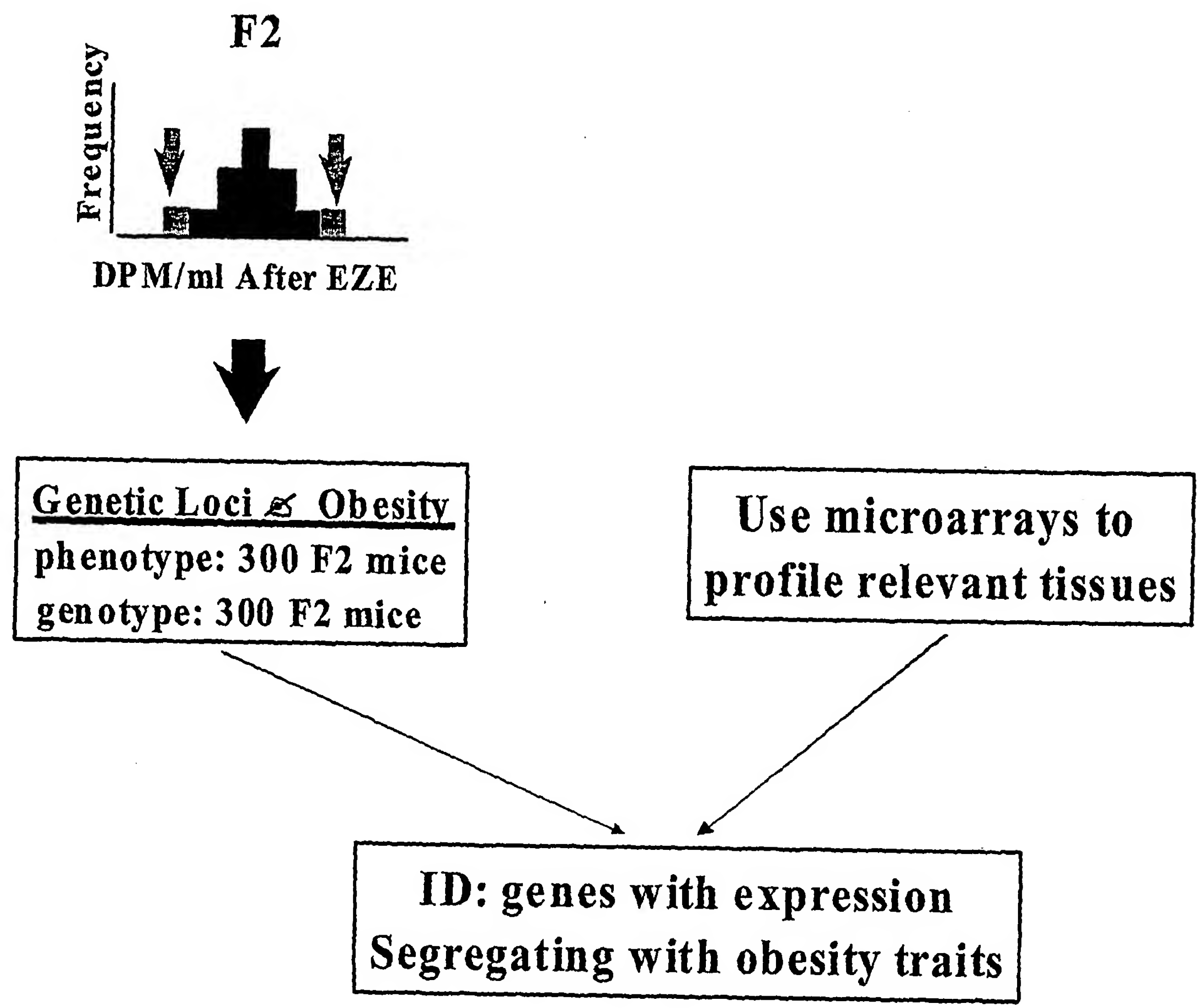

F2
Frequency
DPM/ml After EZE
Genetic Loci ≰ Obesity
phenotype: 300 F2 mice
genotype: 300 F2 mice
Use microarrays to
profile relevant tissues
ID: genes with expression
Segregating with obesity traits

Fig. 56

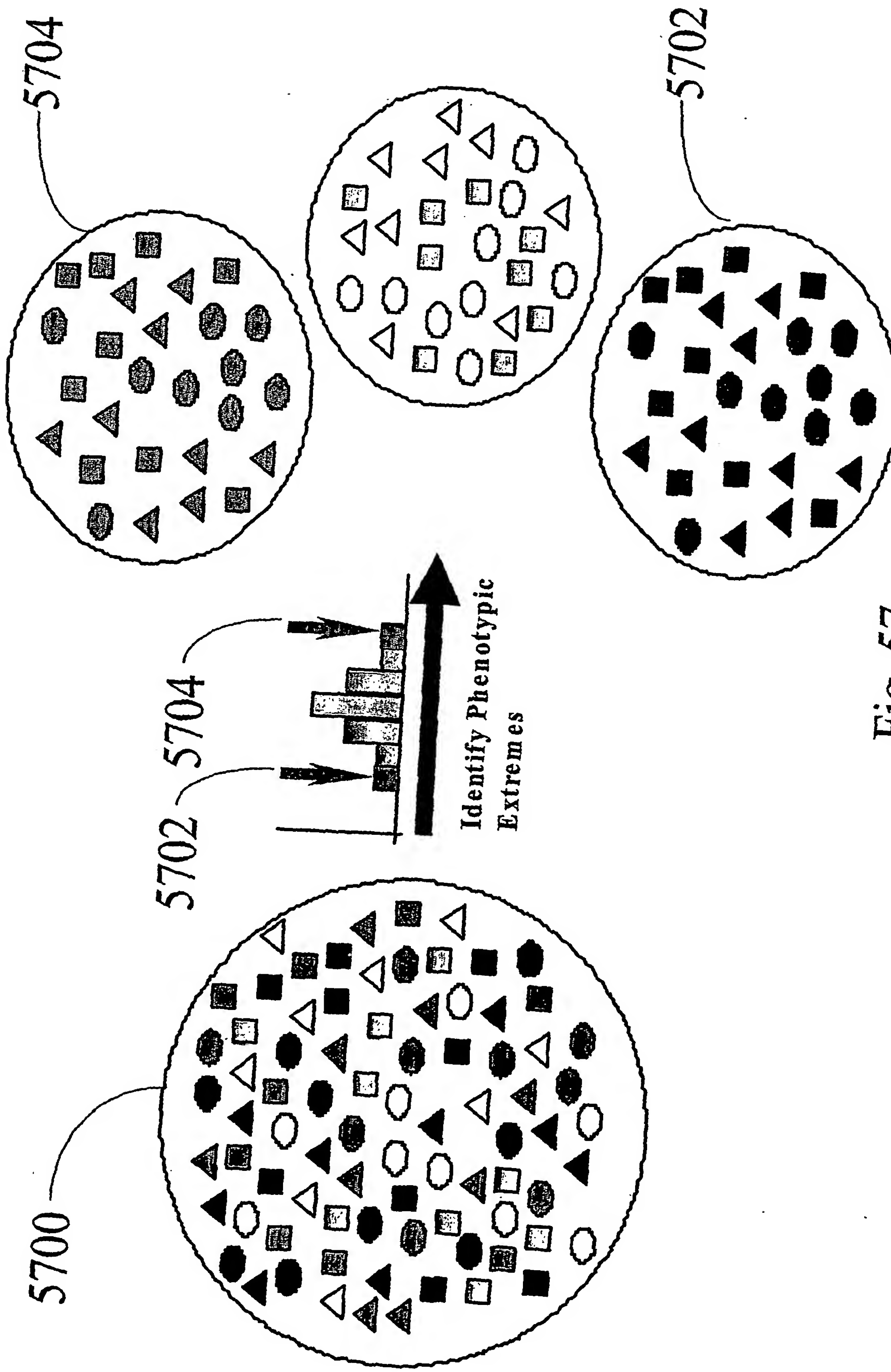



Fig. 57

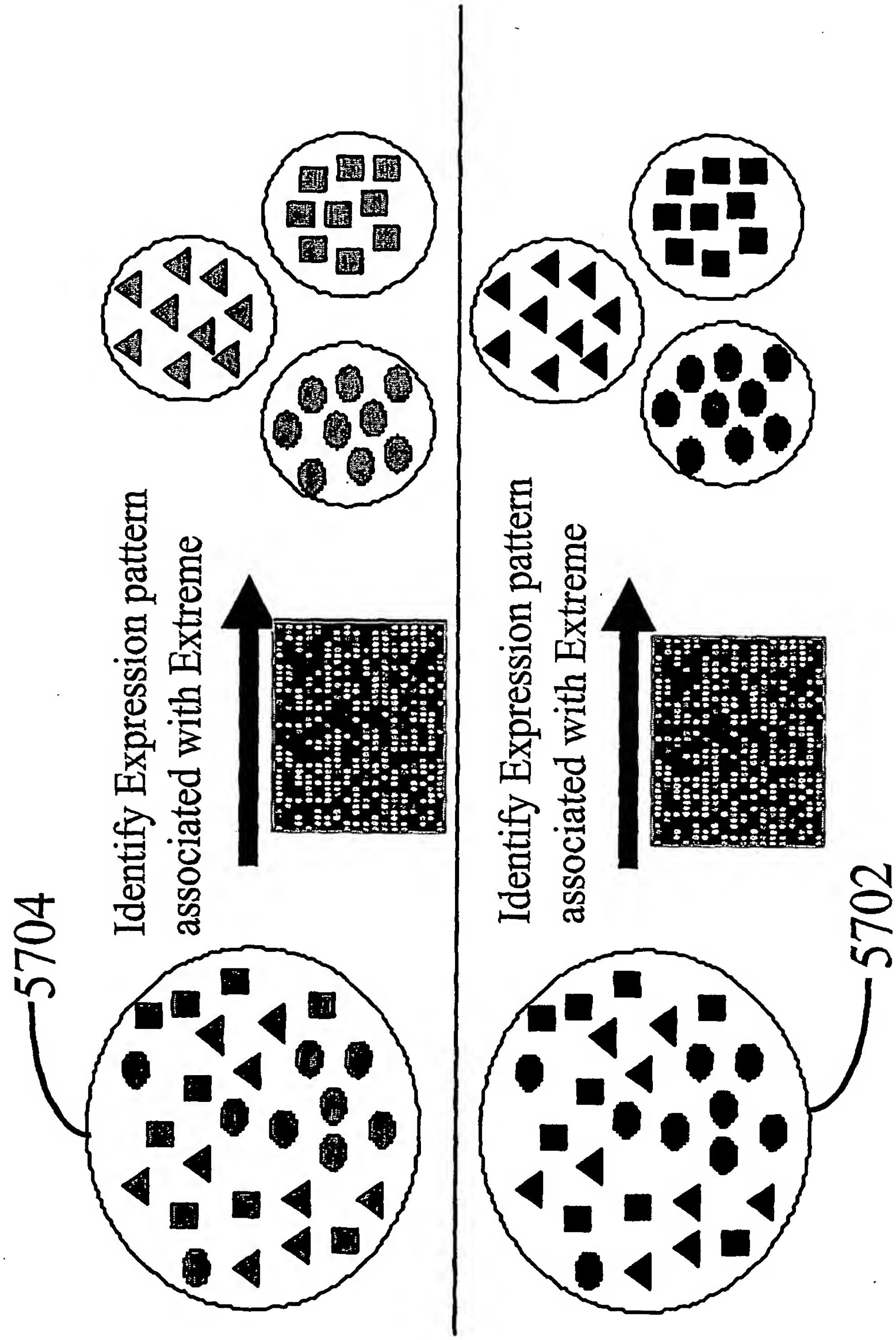

5704
Identify Expression pattern associated with Extreme
5702
Identify Expression pattern associated with Extreme

Fig. 58

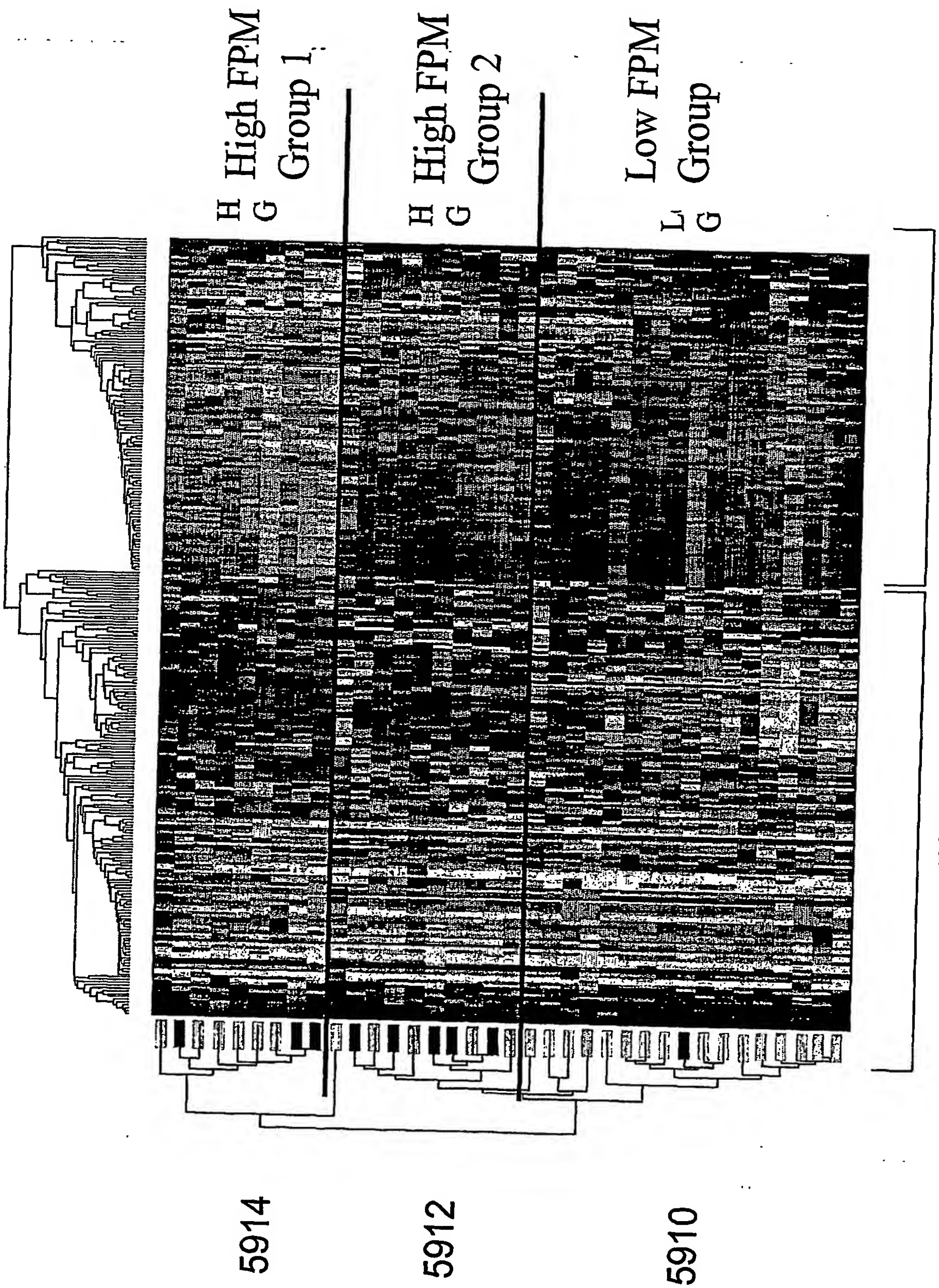

H G High FPM Group 1
H G High FPM Group 2
L G Low FPM Group
5904
5902
5914
5912
5910

Fig. 59

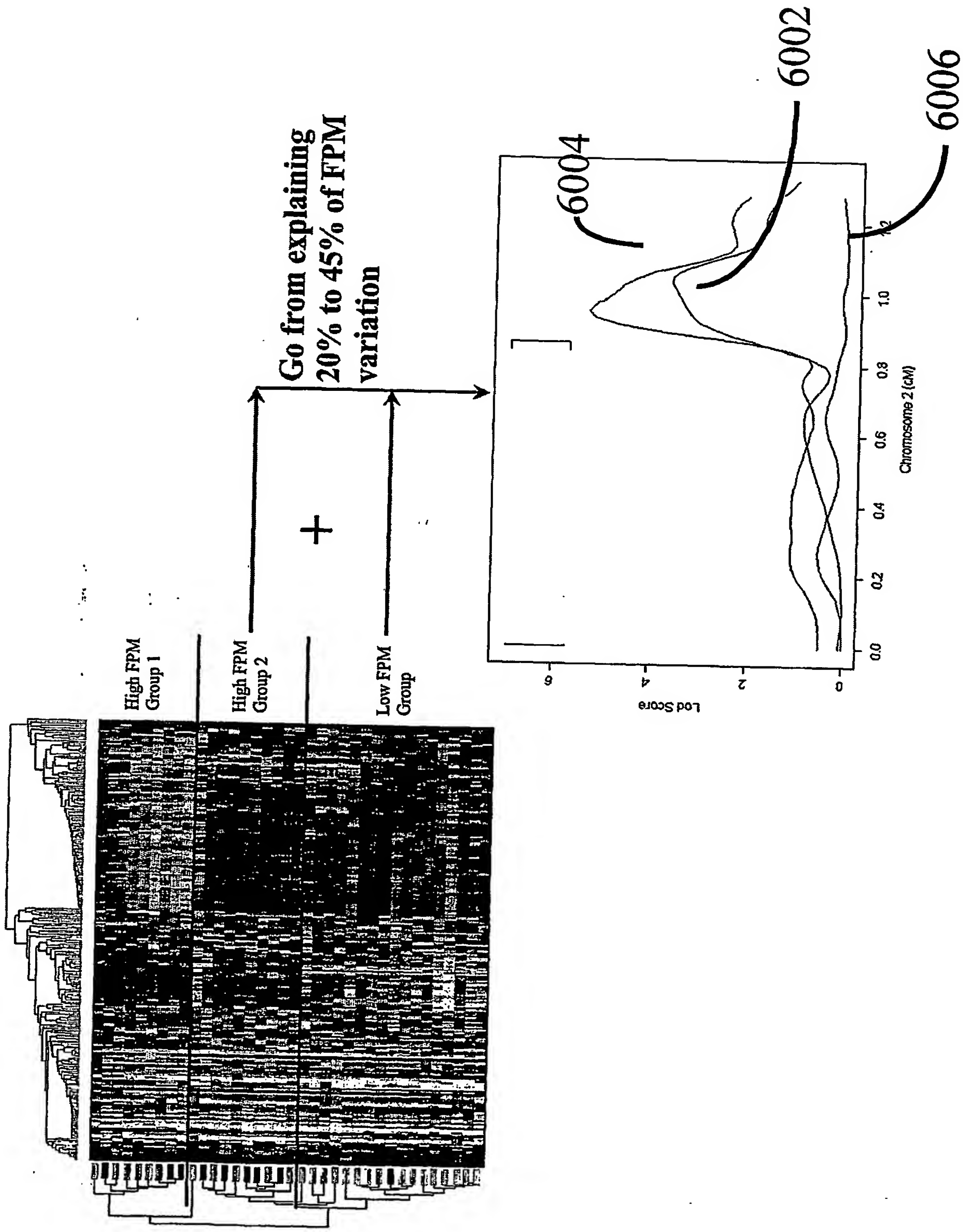

High FPM Group 1
High FPM Group 2
Low FPM Group
+
Go from explaining 20% to 45% of FPM variation
6004
6002
6006
Lod Score
0
2
4
6
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Chromosome 2 (cM)

Fig. 60

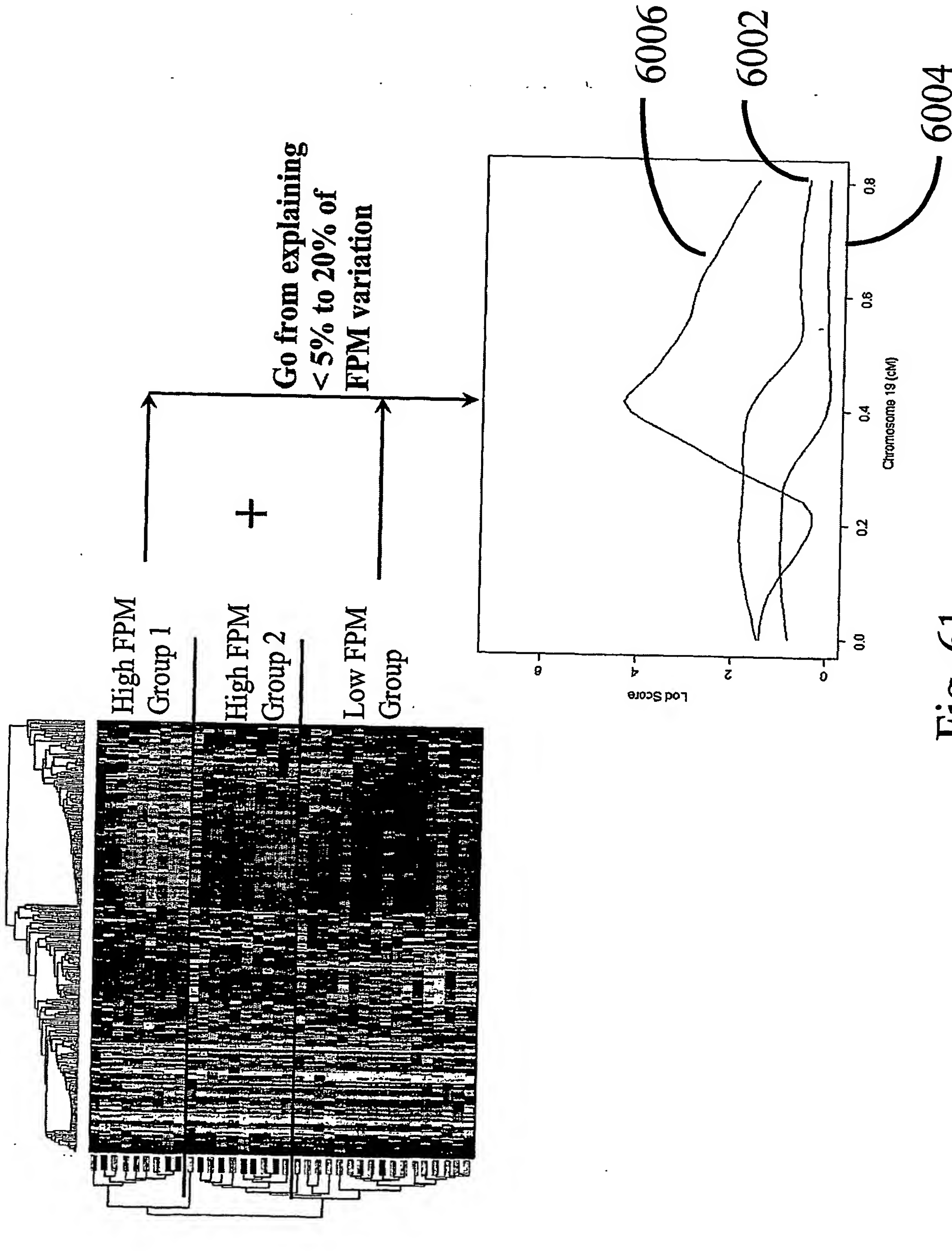

High FPM Group 1
High FPM Group 2
Low FPM Group
+
Go from explaining < 5% to 20% of FPM variation
6006
6002
6004
Lod Score
Chromosome 19 (cM)
0.0
0.2
0.4
0.6
0.8
0
2
4
6

Fig. 61

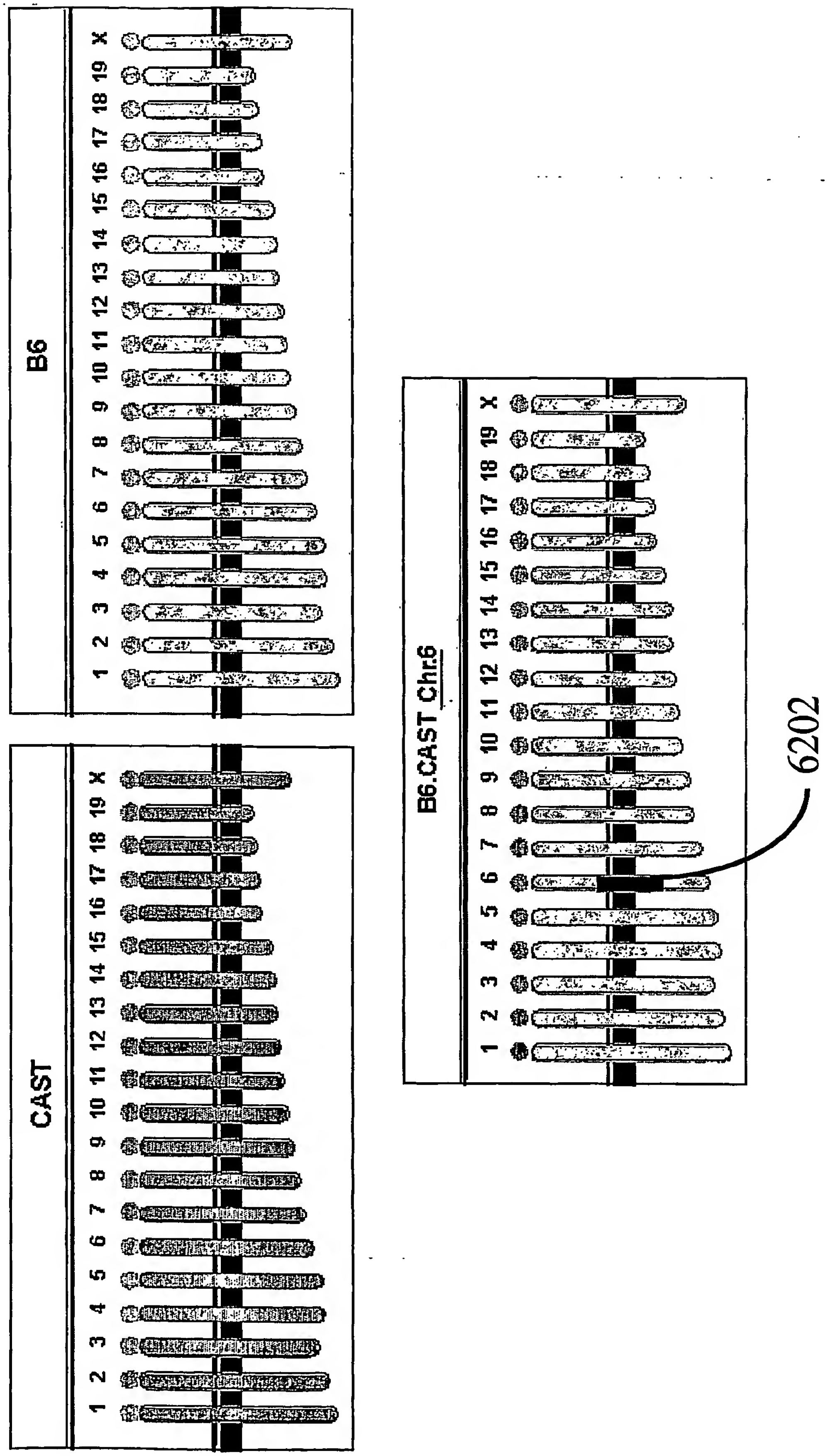

B6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
CAST
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
B6.CAST Chr6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
6202

Fig. 62

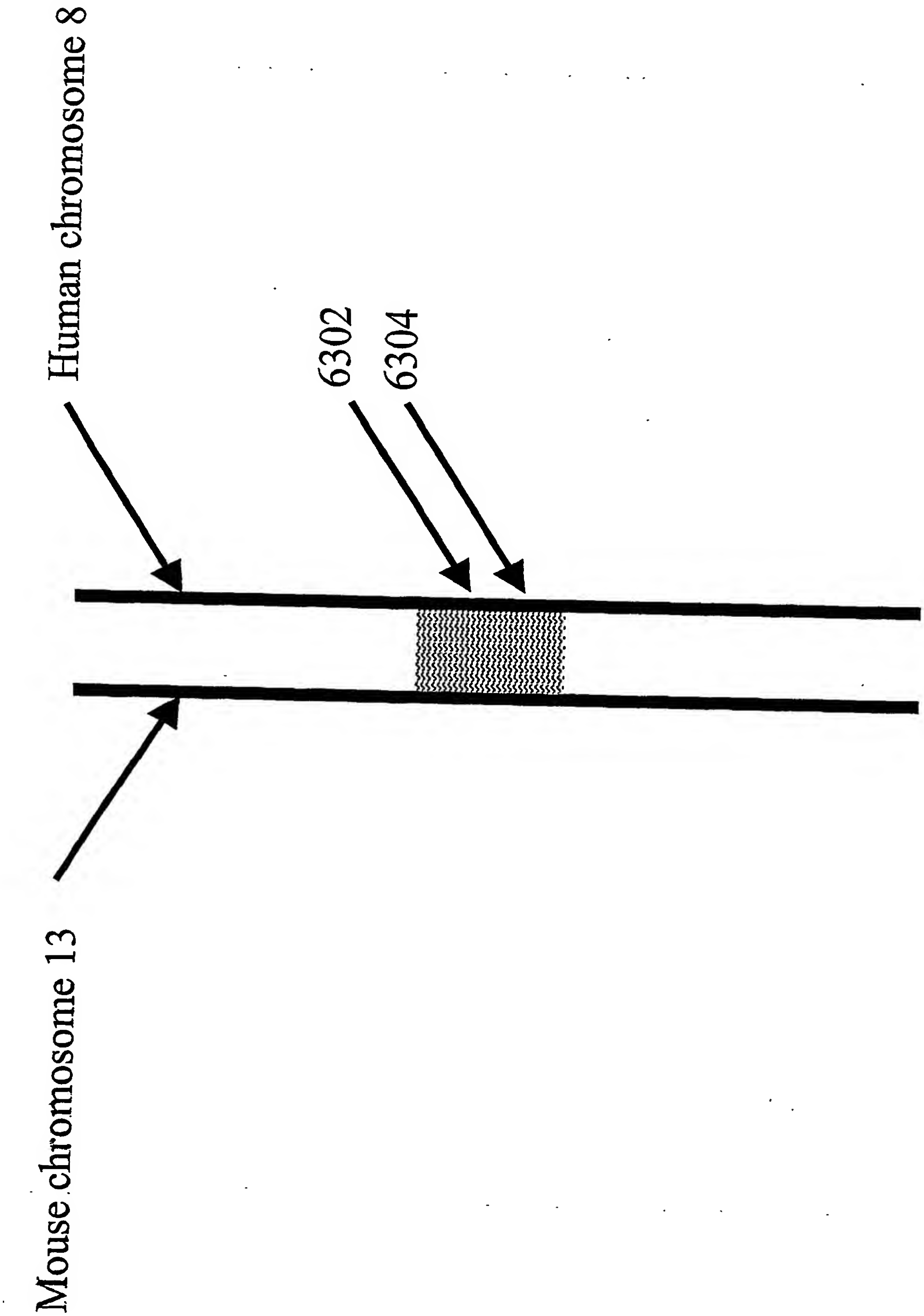



Fig. 63

Lod scores on mouse Chr. 13 (lep, ins, fat, hdl)

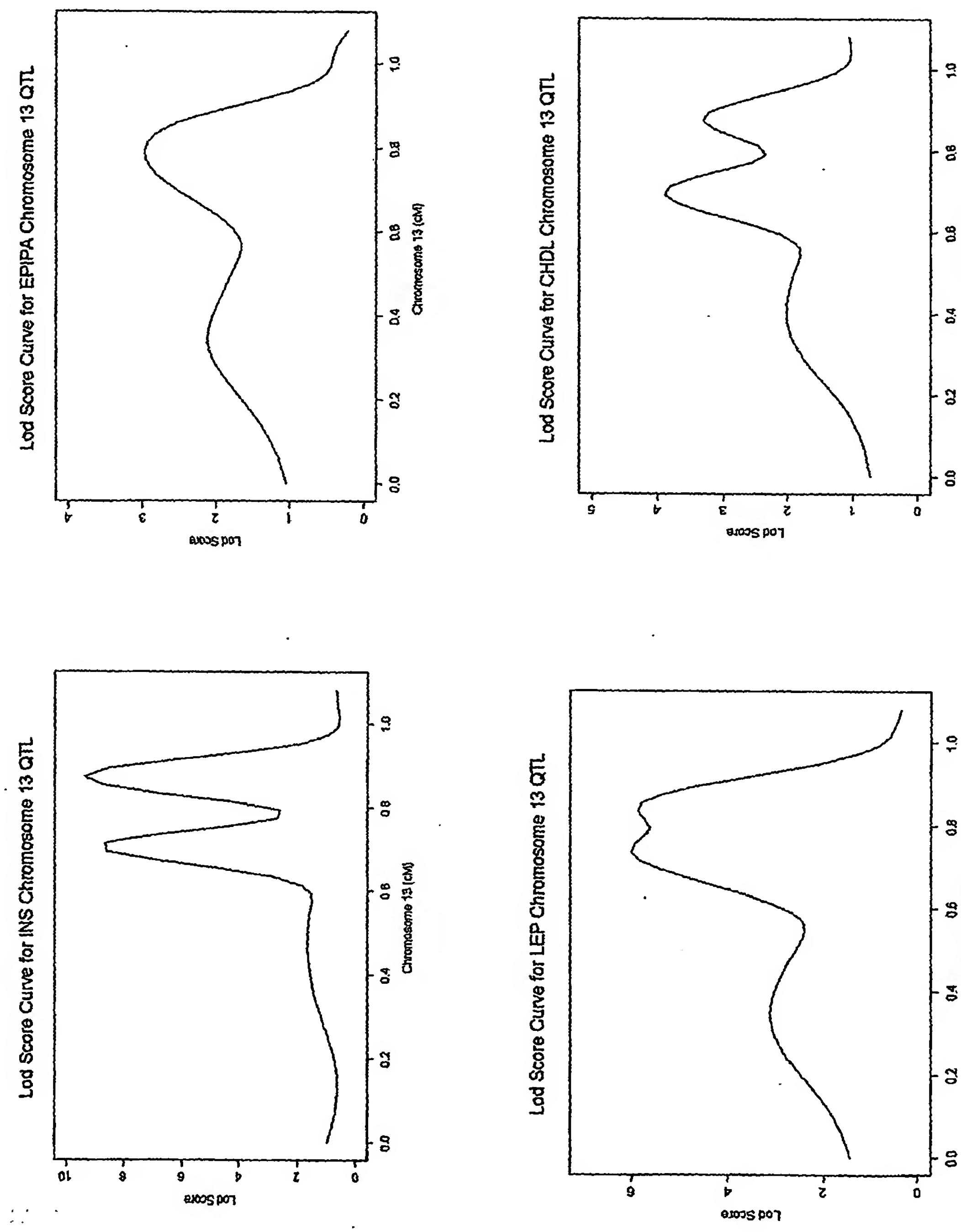



Fig. 64

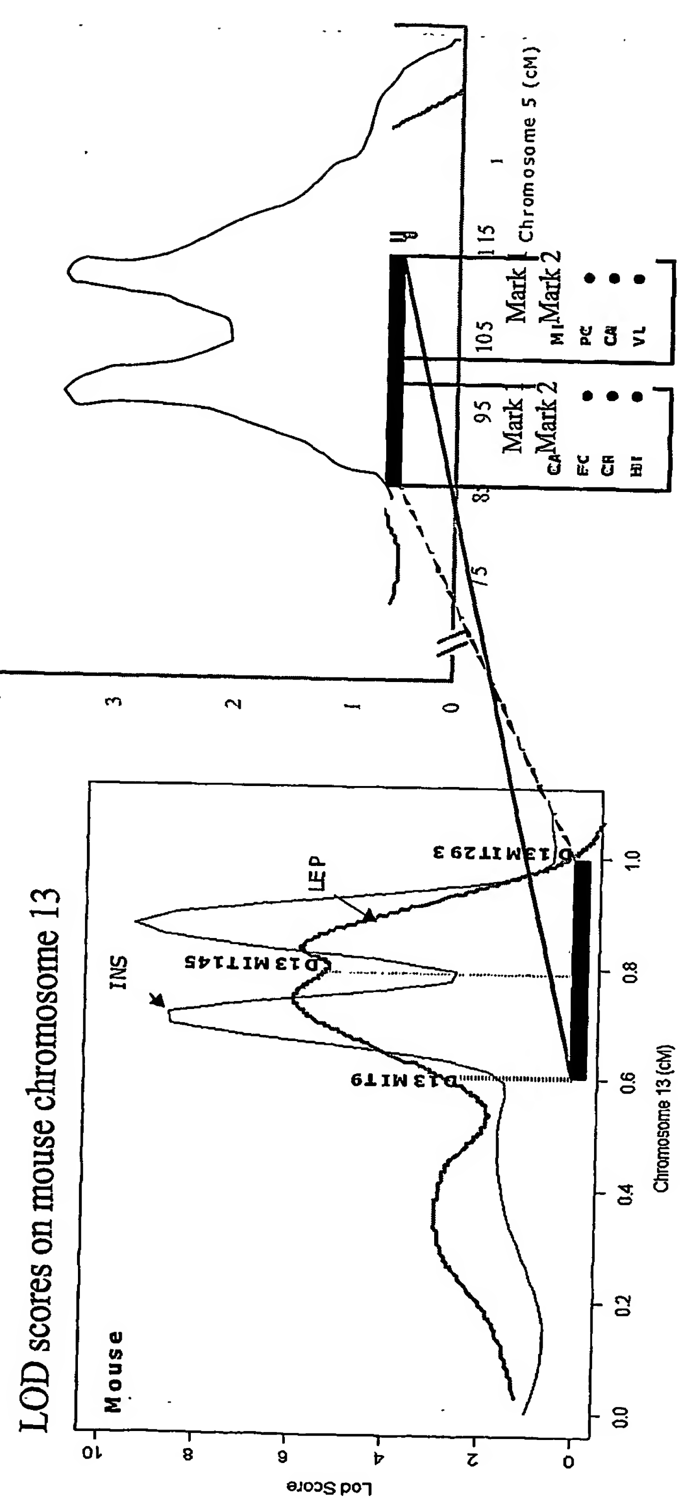
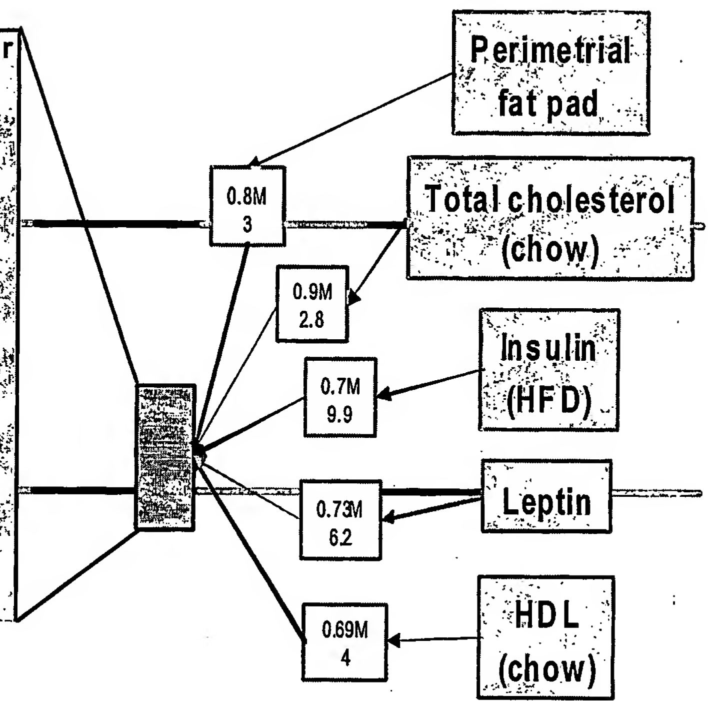

LOD scores on mouse chromosome 13
Mouse
INS
LEP
D13MIT19
D13MIT145
D13MIT293
Lod Score
Chromosome 13 (cM)
LOD scores on human chromosome 8
Chromosome 5 (cM)
Mark 1
Mark 2

Fig. 65

| Name | Chr Location | LOD score | site on Chr 13 (cM) |
|---|---|---|---|
| CCKar | 5 | 20.9 | 87 |
| Foxc2 | 8 | 5.2 | 85 |
| LepR | 4 | 4.1 | 110 |
| DPP IV | 2 | 3.4 | 87 |
| LIPE | 7 | 2.6 | 110 |
| CPT1a | 19 | 2.5 | 110 |
| CB2R | 4 | 2.1 | 110 |
| Orexin | 11 | 2.1 | 110 |
| CTE1 | 12 | 2 | 100 |
| RXRg | 1 | 2 | 100 |
| C330026N13Rik | 13 | 12 | 85 |
| 1810058I14Rik | 13 | 51.1 | 100 |

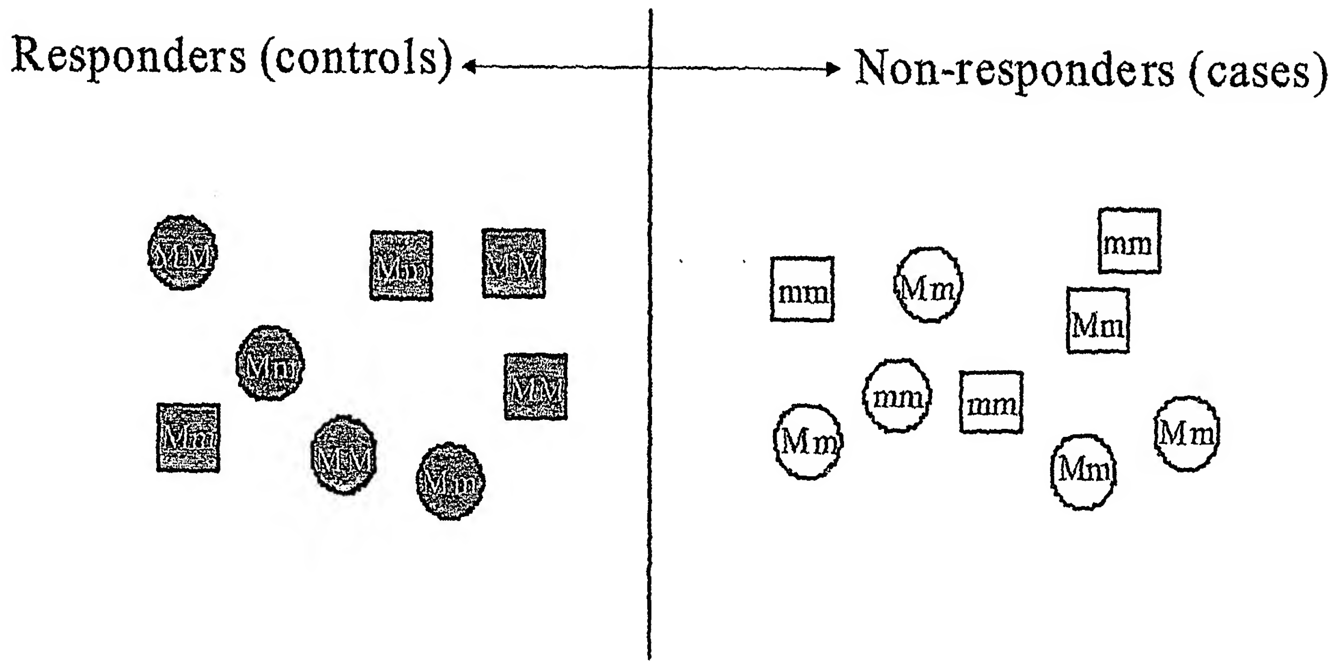



17-19

Fig. 66

# Human Association Analysis

- Given a gene identified in mouse for EZE response, we can directly test whether polymorphisms in this gene in human populations are associated with this trait




Ex: Is the frequency of the polymorphism equal between EZE responders and non-responders?

Fig. 67

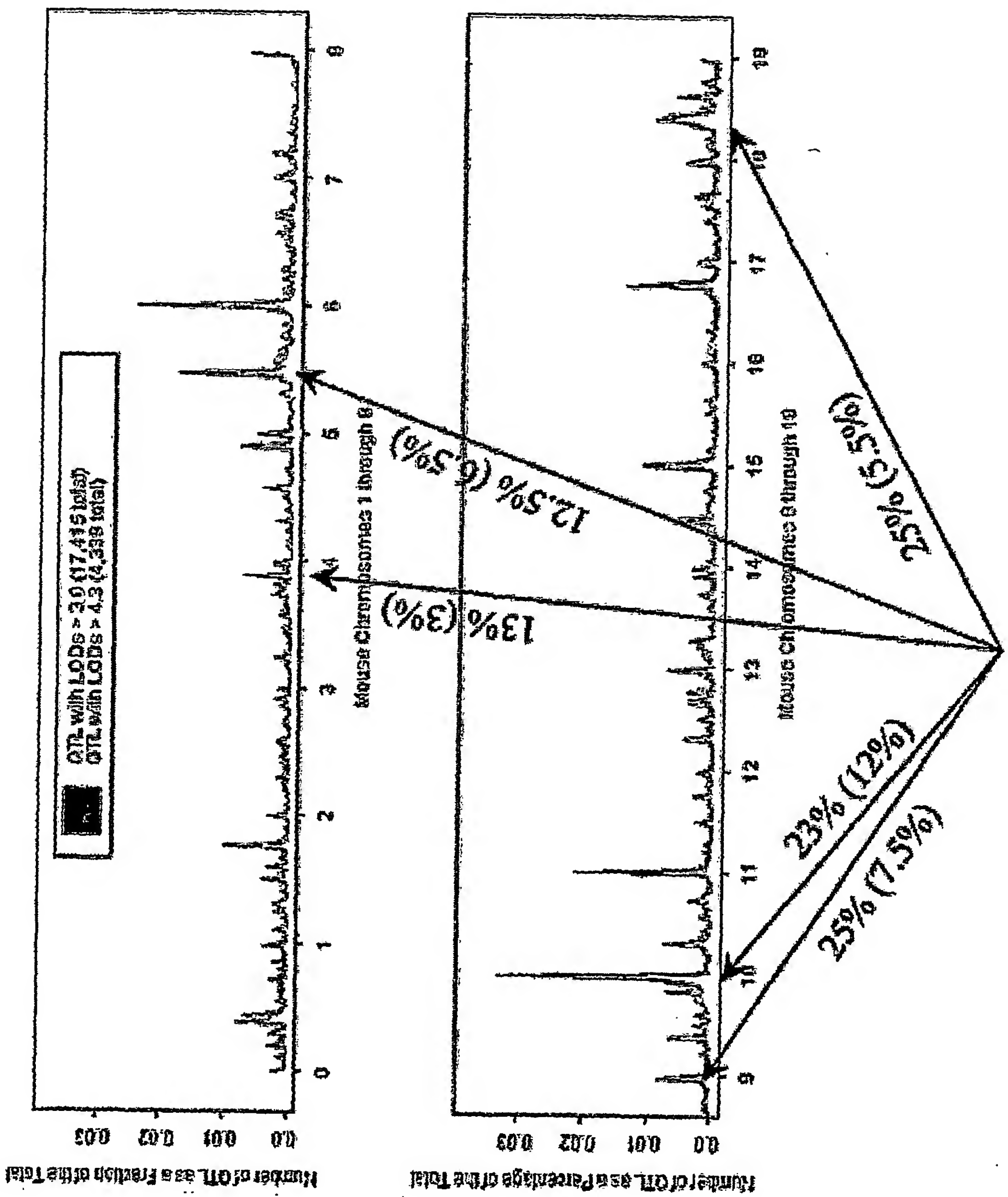

QTL with LODs > 3.0 (17,415 total)
QTL with LODs > 4.3 (4,338 total)
Number of QTL as a Fraction of the Total
Mouse Chromosomes 1 through 8
Number of QTL as a Percentage of the Total
Mouse Chromosomes 9 through 19
13% (3%)
12.5% (6.5%)
25% (5.5%)
23% (12%)
25% (7.5%)

Fig. 68

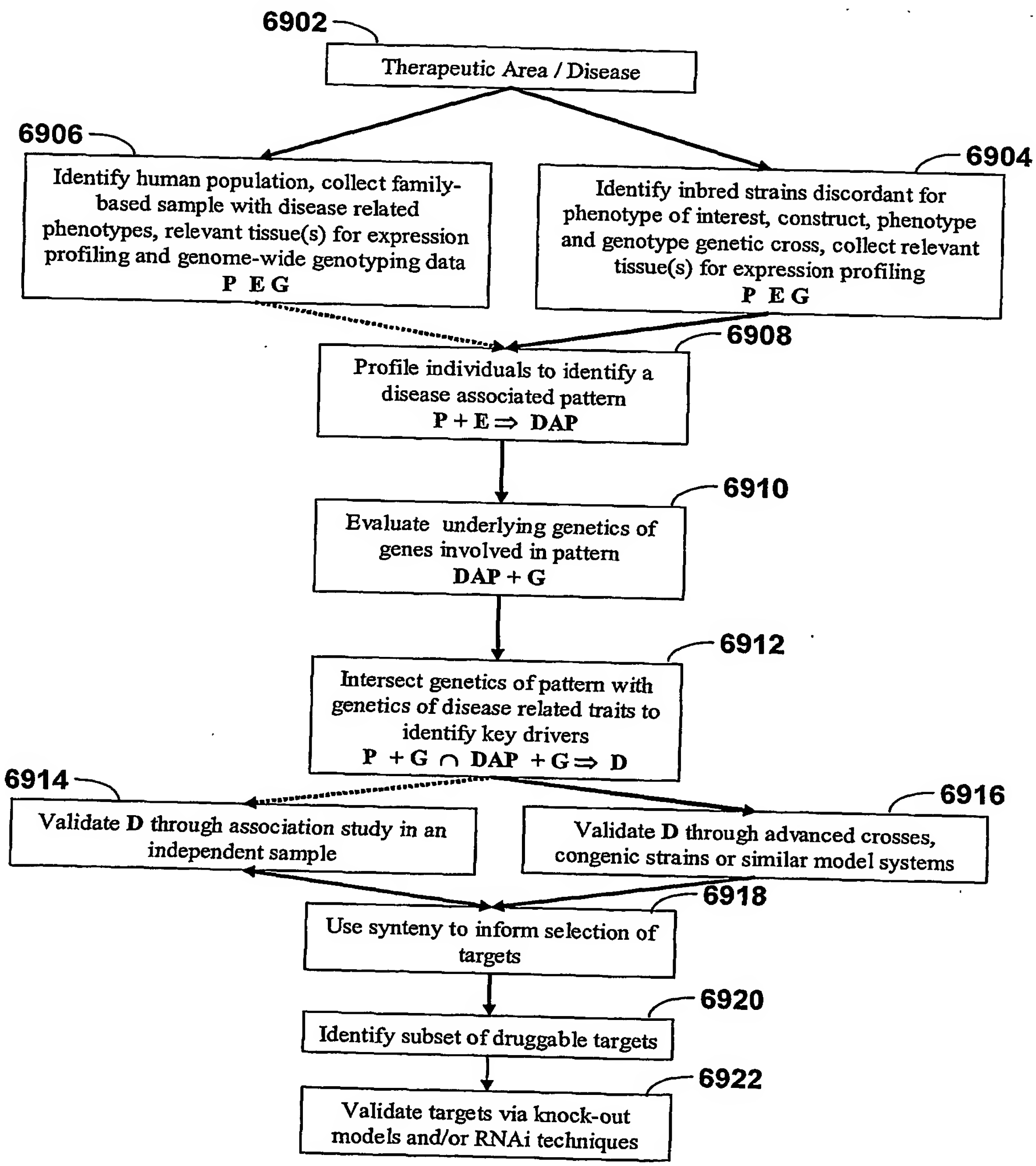

6902
Therapeutic Area / Disease
6906
Identify human population, collect family-based sample with disease related phenotypes, relevant tissue(s) for expression profiling and genome-wide genotyping data
P E G
6904
Identify inbred strains discordant for phenotype of interest, construct, phenotype and genotype genetic cross, collect relevant tissue(s) for expression profiling
P E G
6908
Profile individuals to identify a disease associated pattern
P + E ⇒ DAP
6910
Evaluate underlying genetics of genes involved in pattern
DAP + G
6912
Intersect genetics of pattern with genetics of disease related traits to identify key drivers
P + G ∩ DAP + G ⇒ D
6914
Validate D through association study in an independent sample
6916
Validate D through advanced crosses, congenic strains or similar model systems
6918
Use synteny to inform selection of targets
6920
Identify subset of druggable targets
6922
Validate targets via knock-out models and/or RNAi techniques

Fig. 69

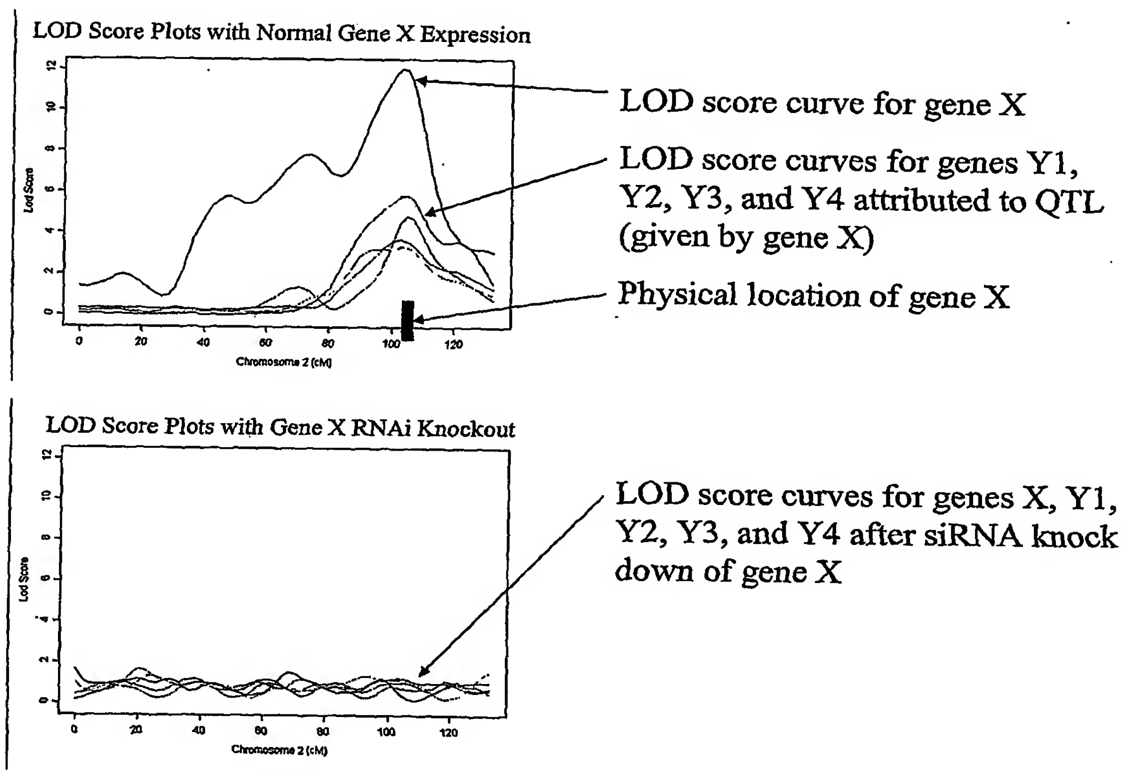

LOD Score Plots with Normal Gene X Expression
LOD score curve for gene X
LOD score curves for genes Y1, Y2, Y3, and Y4 attributed to QTL (given by gene X)
Physical location of gene X
Lod Score
Chromosome 2 (cM)
LOD Score Plots with Gene X RNAi Knockout
LOD score curves for genes X, Y1, Y2, Y3, and Y4 after siRNA knock down of gene X
Lod Score
Chromosome 2 (cM)

Fig. 70

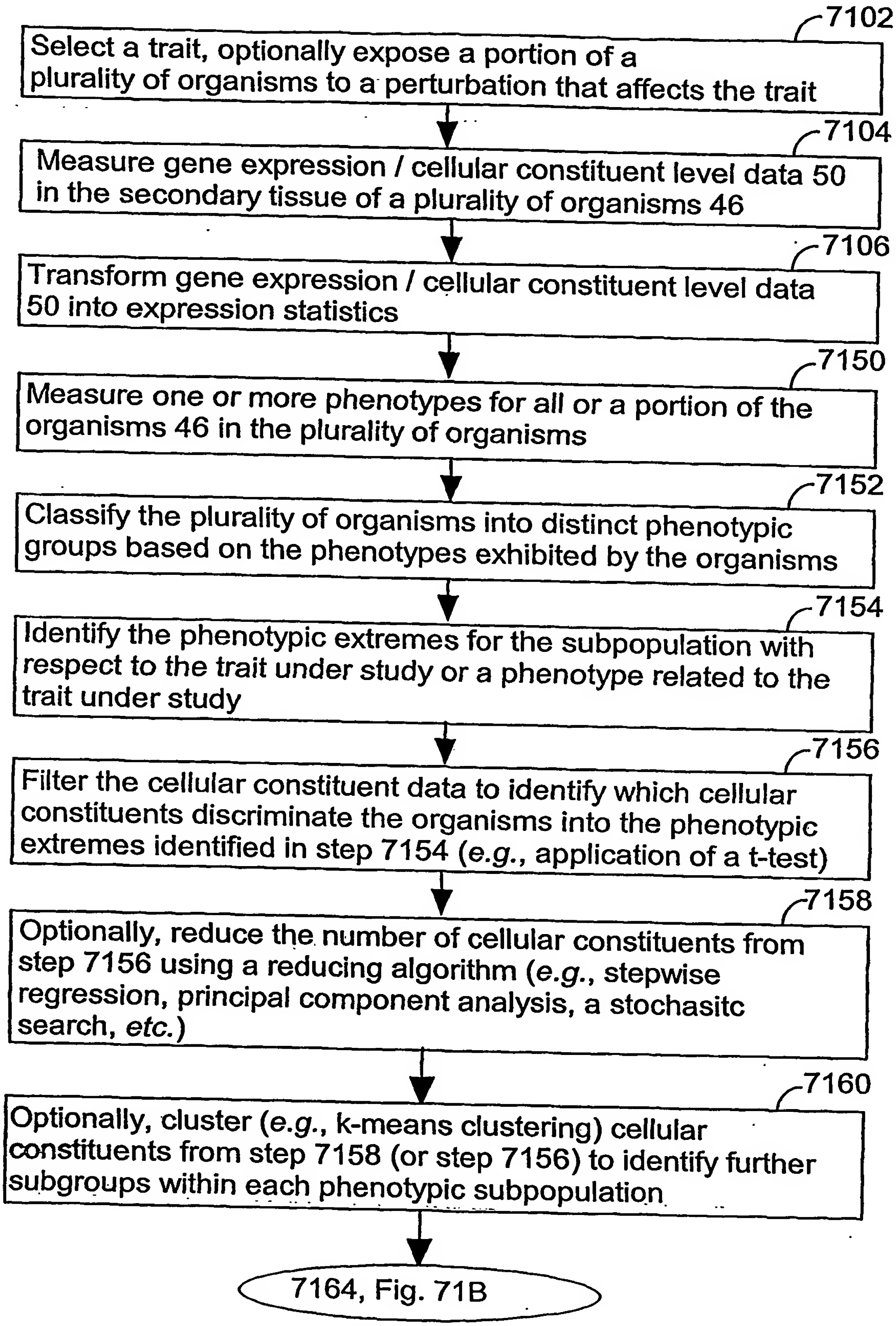

7102
Select a trait, optionally expose a portion of a plurality of organisms to a perturbation that affects the trait
7104
Measure gene expression / cellular constituent level data 50 in the secondary tissue of a plurality of organisms 46
7106
Transform gene expression / cellular constituent level data 50 into expression statistics
7150
Measure one or more phenotypes for all or a portion of the organisms 46 in the plurality of organisms
7152
Classify the plurality of organisms into distinct phenotypic groups based on the phenotypes exhibited by the organisms
7154
Identify the phenotypic extremes for the subpopulation with respect to the trait under study or a phenotype related to the trait under study
7156
Filter the cellular constituent data to identify which cellular constituents discriminate the organisms into the phenotypic extremes identified in step 7154 (e.g., application of a t-test)
7158
Optionally, reduce the number of cellular constituents from step 7156 using a reducing algorithm (e.g., stepwise regression, principal component analysis, a stochasitc search, etc.)
7160
Optionally, cluster (e.g., k-means clustering) cellular constituents from step 7158 (or step 7156) to identify further subgroups within each phenotypic subpopulation
7164, Fig. 71B

**FIG. 71A**

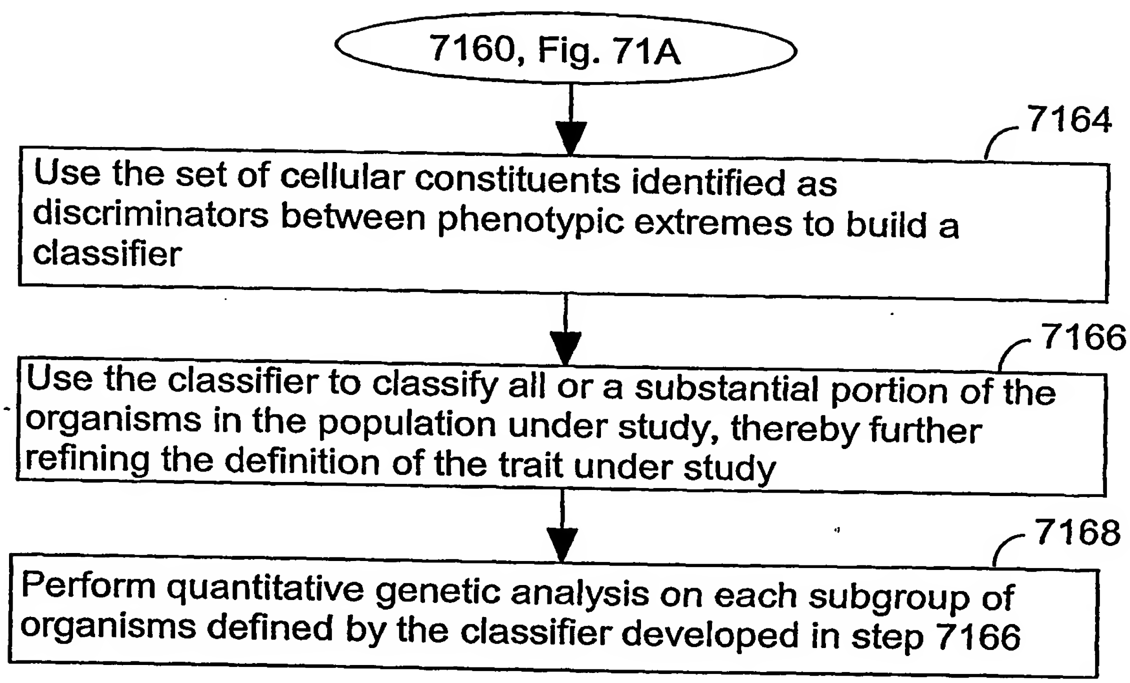

7160, Fig. 71A
7164
Use the set of cellular constituents identified as discriminators between phenotypic extremes to build a classifier
7166
Use the classifier to classify all or a substantial portion of the organisms in the population under study, thereby further refining the definition of the trait under study
7168
Perform quantitative genetic analysis on each subgroup of organisms defined by the classifier developed in step 7166

FIG. 71B

| | Phenotype 1 | • • • | Phenotype M | CC 48-1 | • • • | CC 48-Z |
|---|---|---|---|---|---|---|
| Organism 46-1 | Amount 7201-1-1 | • • • | Amount 7201-1-M | Level 50-1-1 | • • • | Level 50-1-Z |
| Organism 46-2 | Amount 7201-2-1 | • • • | Amount 7201-2-M | Level 50-2-1 | • • • | Level 50-2-Z |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Organism 46-N | Amount 7201-N-1 | • • • | Amount 7201-N-M | Level 50-N-1 | • • • | Level 50-N-Z |

FIG. 72

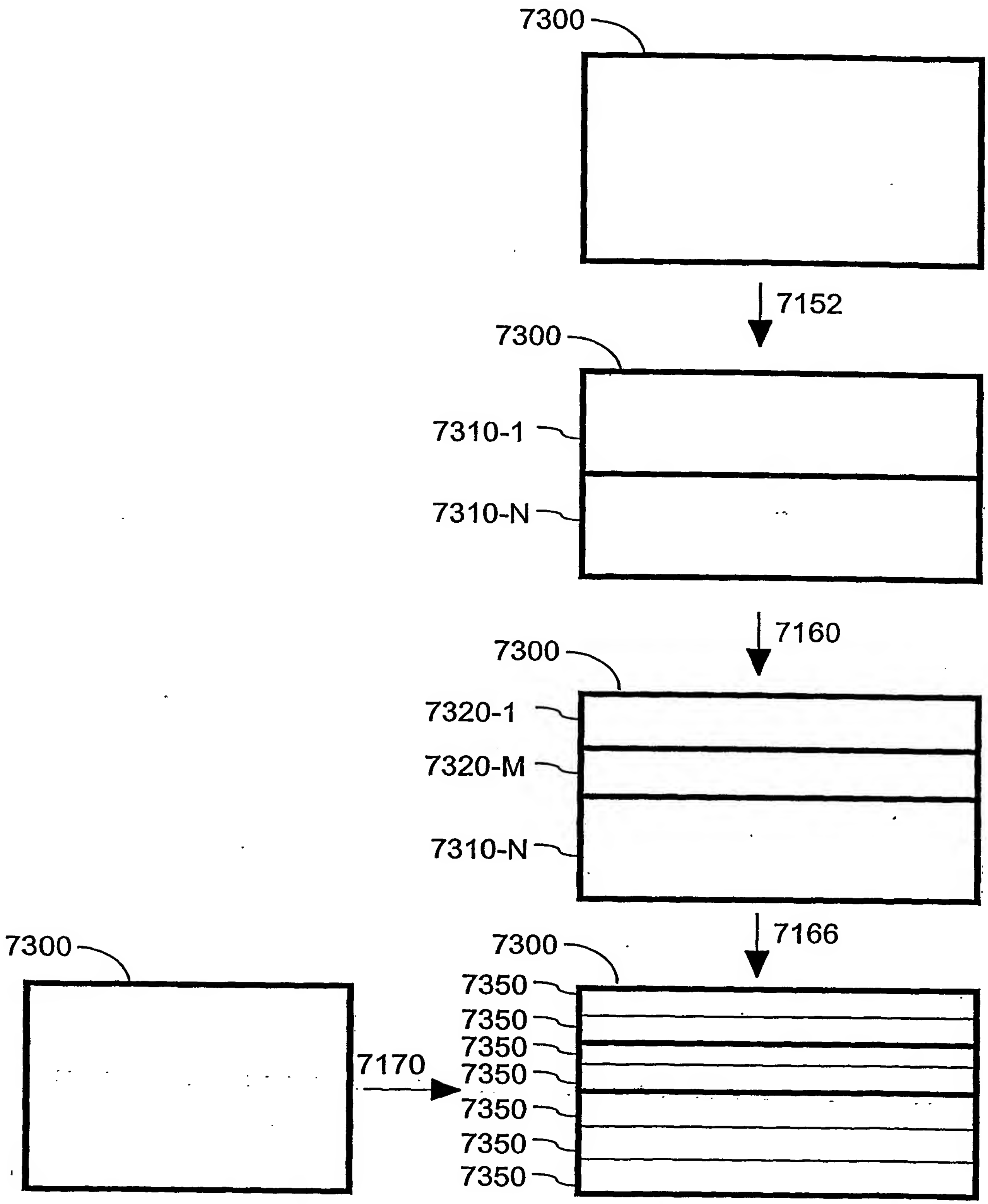

7300
7152
7300
7310-1
7310-N
7160
7300
7320-1
7320-M
7310-N
7166
7300
7300
7350
7350
7350
7350
7350
7350
7350
7170

FIG. 73

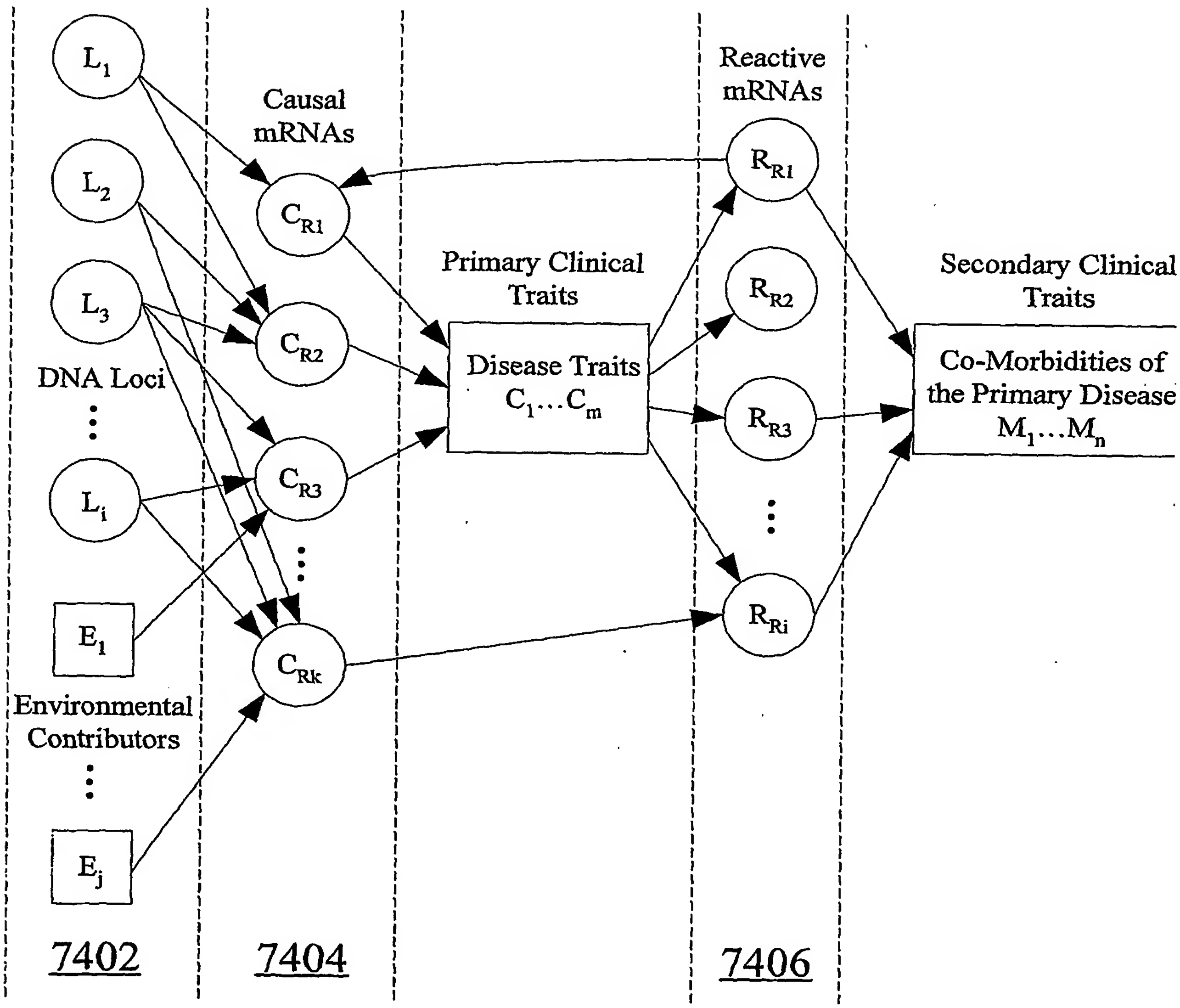

Causal mRNAs
Reactive mRNAs
$L_1$
$L_2$
$L_3$
DNA Loci
$L_i$
$E_1$
Environmental Contributors
$E_j$
$C_{R1}$
$C_{R2}$
$C_{R3}$
$C_{Rk}$
Primary Clinical Traits
Disease Traits $C_1 \ldots C_m$
$R_{R1}$
$R_{R2}$
$R_{R3}$
$R_{Ri}$
Secondary Clinical Traits
Co-Morbidities of the Primary Disease $M_1 \ldots M_n$
7402
7404
7406

Fig. 74

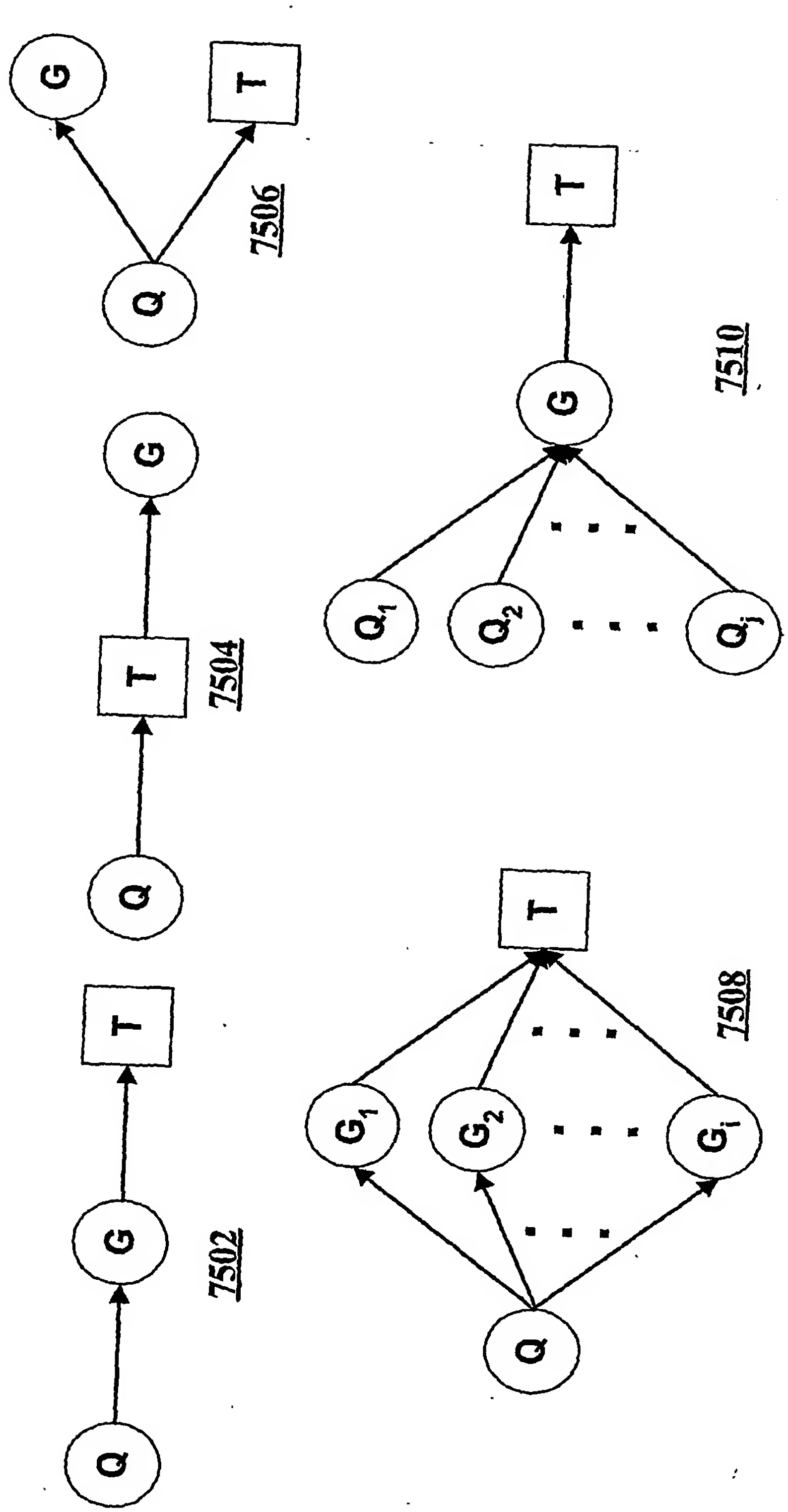



Fig. 75A

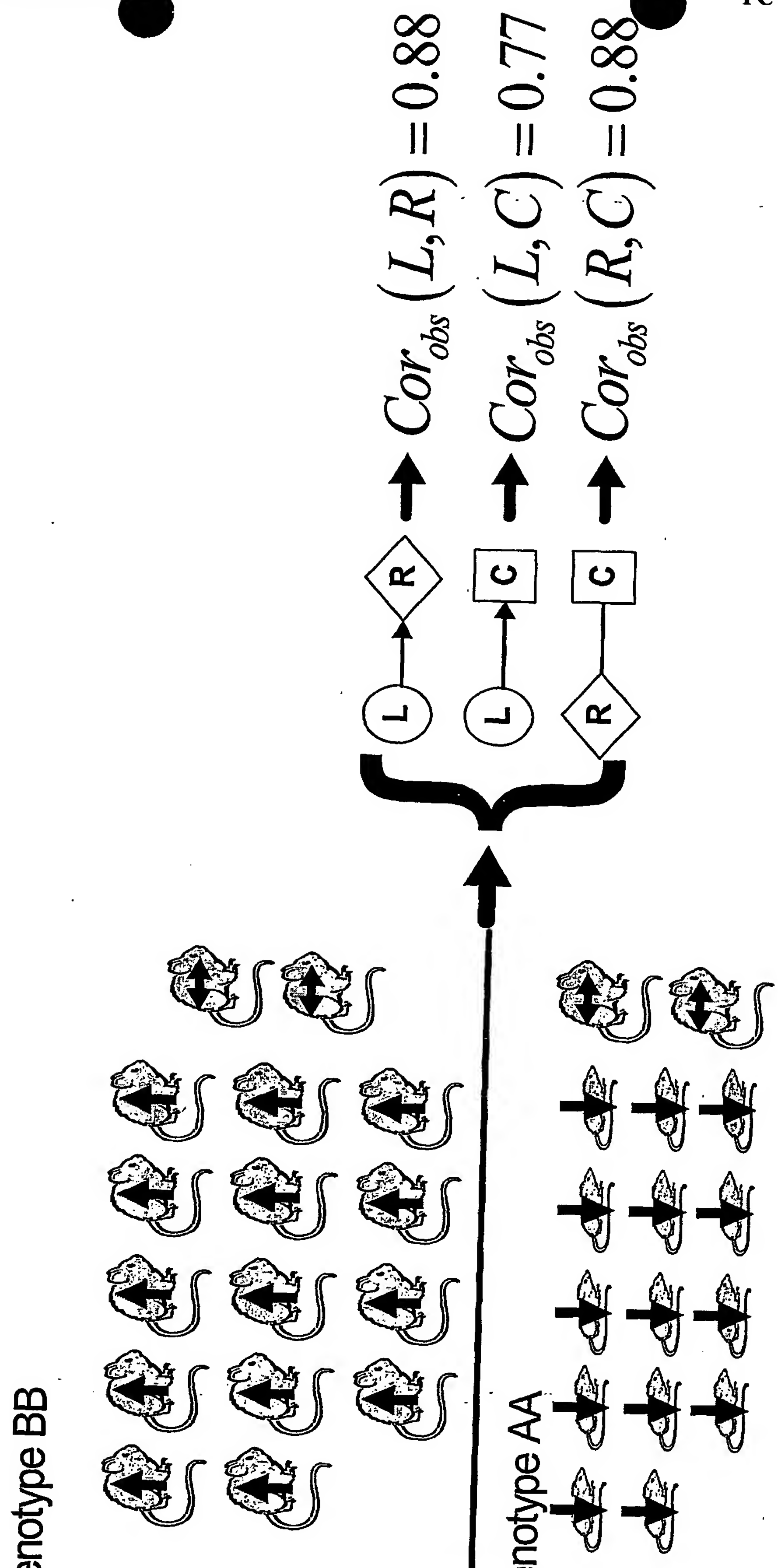

Genotype BB
Genotype AA
L
R
L
C
R
C
$Cor_{obs}(L,R)=0.88$
$Cor_{obs}(L,C)=0.77$
$Cor_{obs}(R,C)=0.88$

Figure 75B

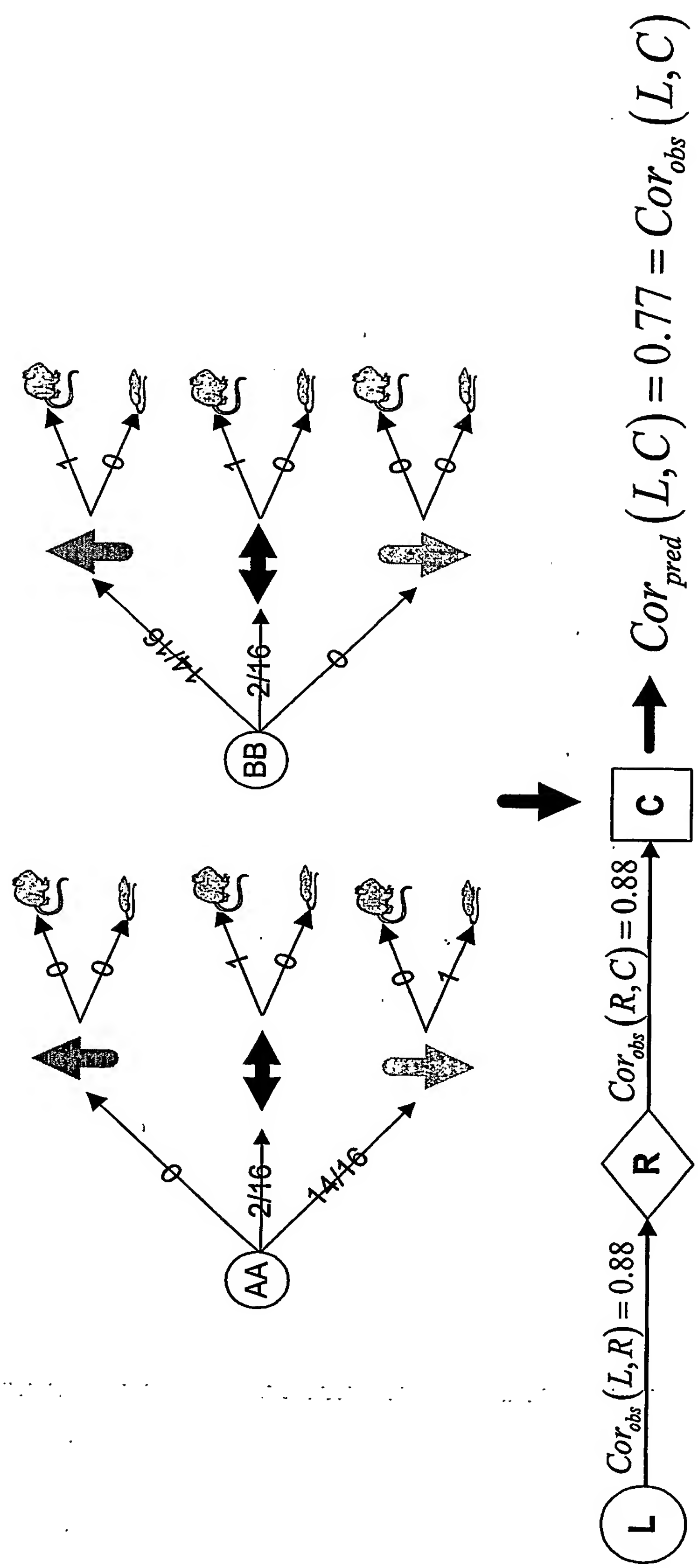

AA
BB
2/46
14/46
0
1
L
R
C
$Cor_{obs}(L,R)=0.88$
$Cor_{obs}(R,C)=0.88$
$Cor_{pred}(L,C)=0.77=Cor_{obs}(L,C)$

Figure 75C

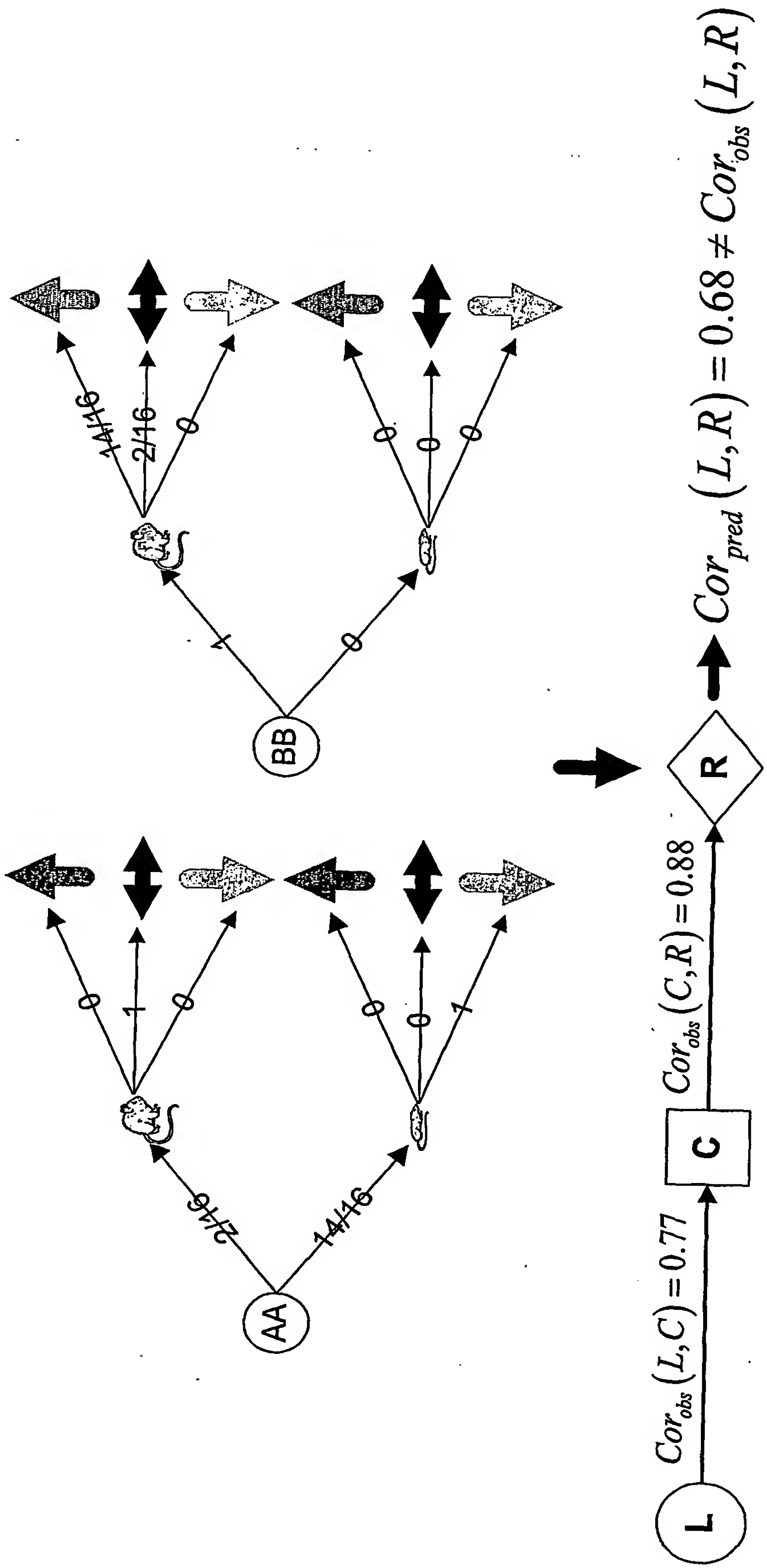

AA
2/16
14/16
0
1
0
0
0
1
BB
1
0
14/16
2/16
0
0
0
0
L
$Cor_{obs}(L,C)=0.77$
C
$Cor_{obs}(C,R)=0.88$
R
$Cor_{pred}(L,R)=0.68 \neq Cor_{obs}(L,R)$

Figure 75D

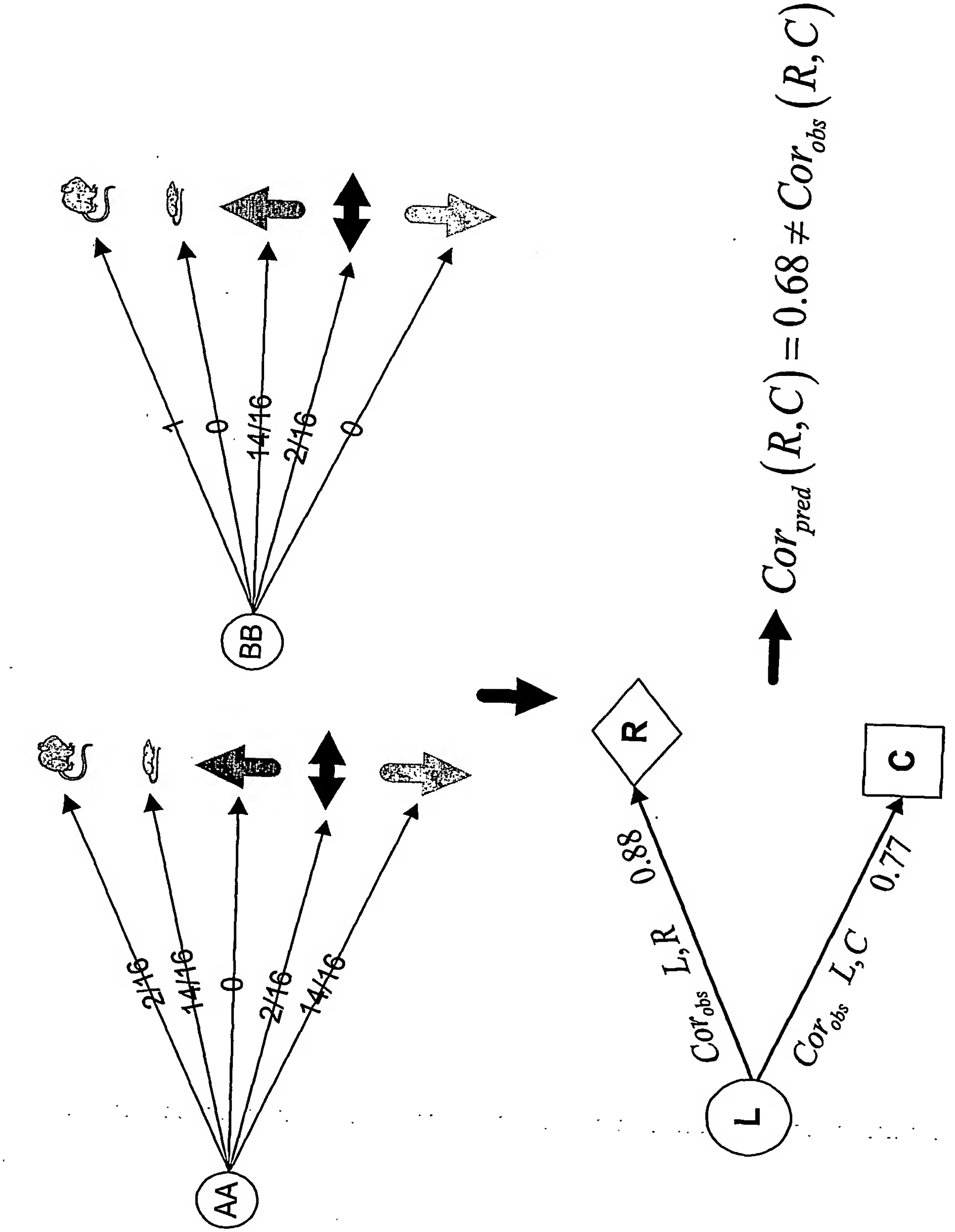

AA
2/16
14/16
0
2/16
14/16
BB
1
0
14/16
2/16
0
L
R
C
$Cor_{obs}\ L,R\ 0.88$
$Cor_{obs}\ L,C\ 0.77$
$Cor_{pred}(R,C) = 0.68 \neq Cor_{obs}(R,C)$

Figure 75E

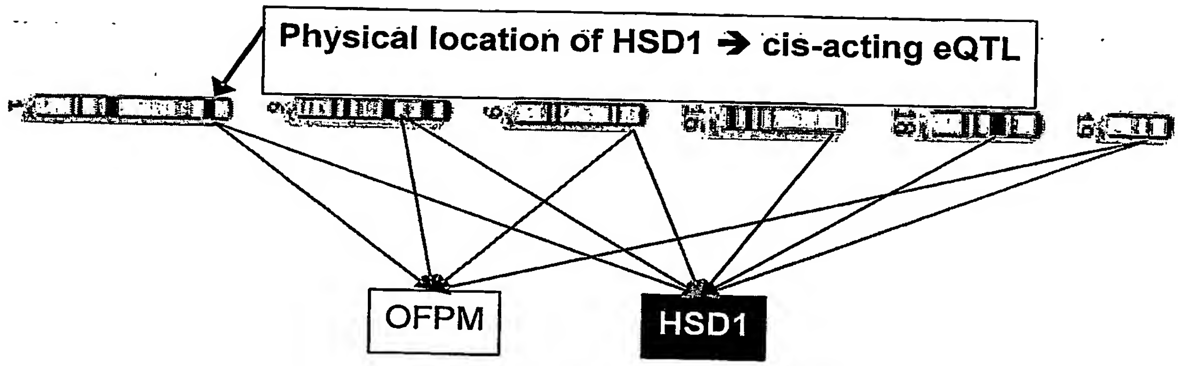

Physical location of HSD1 → cis-acting eQTL
OFPM
HSD1

Fig. 76

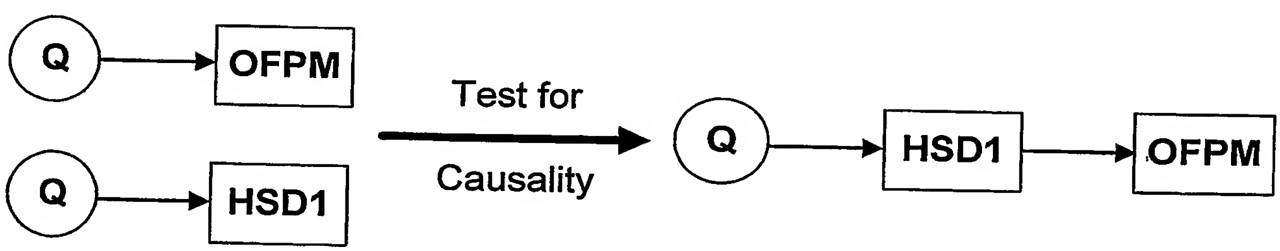
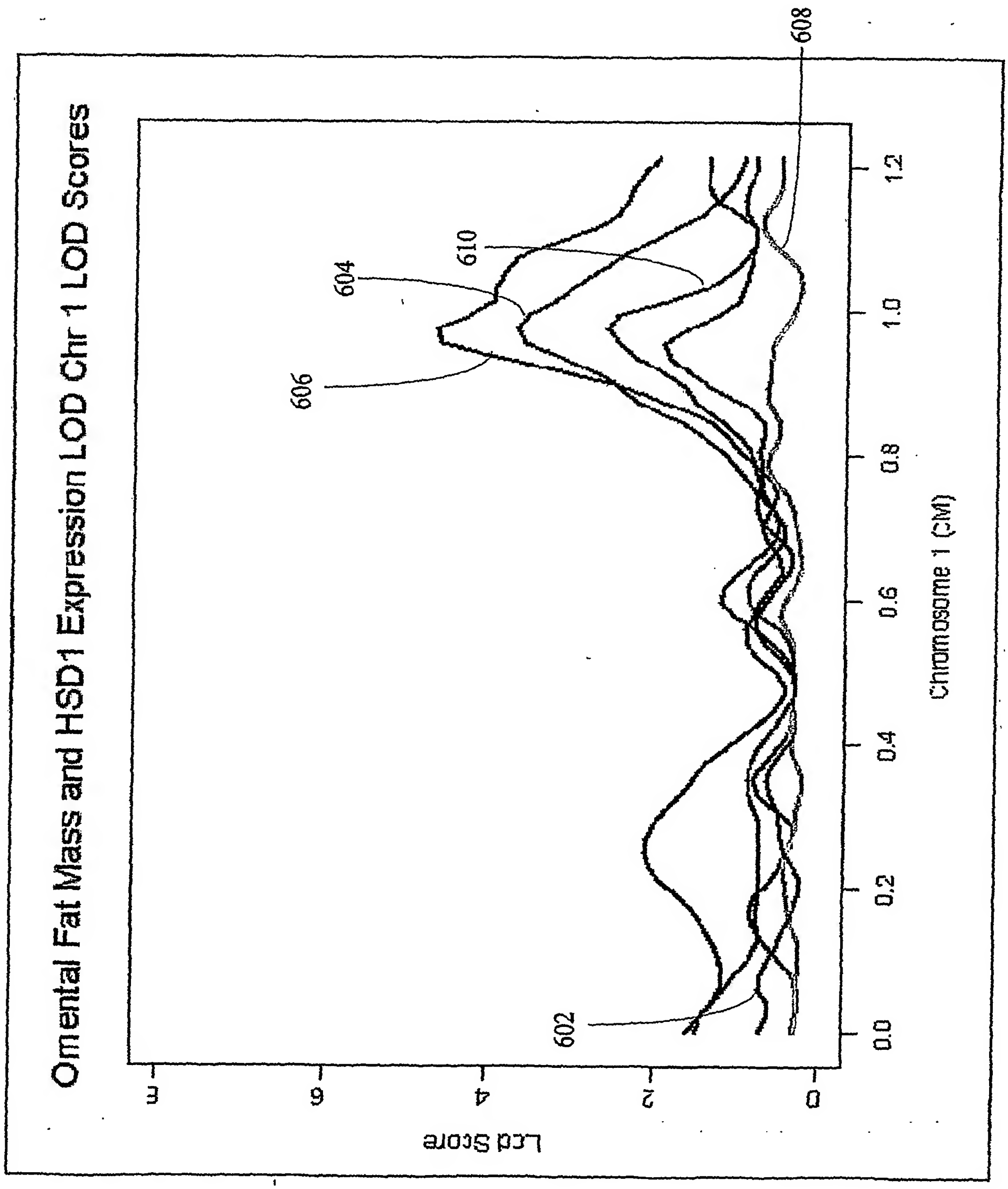

Q
OFPM
Q
HSD1
Test for
Causality
Q
HSD1
OFPM

Fig. 77




Chromosome 6 (cM)

Fig. 78

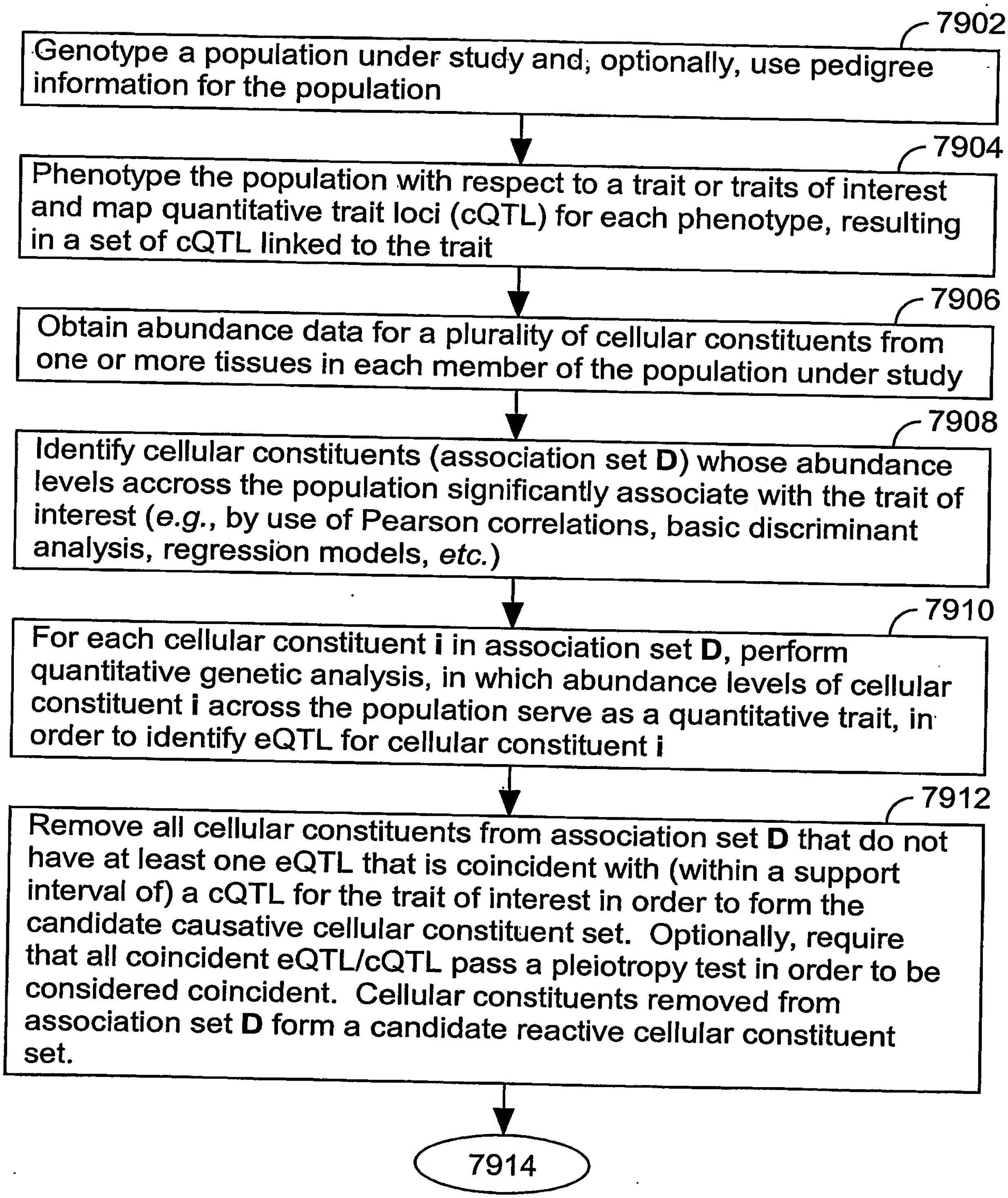

7902
Genotype a population under study and, optionally, use pedigree information for the population
7904
Phenotype the population with respect to a trait or traits of interest and map quantitative trait loci (cQTL) for each phenotype, resulting in a set of cQTL linked to the trait
7906
Obtain abundance data for a plurality of cellular constituents from one or more tissues in each member of the population under study
7908
Identify cellular constituents (association set **D**) whose abundance levels accross the population significantly associate with the trait of interest (*e.g.*, by use of Pearson correlations, basic discriminant analysis, regression models, *etc.*)
7910
For each cellular constituent **i** in association set **D**, perform quantitative genetic analysis, in which abundance levels of cellular constituent **i** across the population serve as a quantitative trait, in order to identify eQTL for cellular constituent **i**
7912
Remove all cellular constituents from association set **D** that do not have at least one eQTL that is coincident with (within a support interval of) a cQTL for the trait of interest in order to form the candidate causative cellular constituent set. Optionally, require that all coincident eQTL/cQTL pass a pleiotropy test in order to be considered coincident. Cellular constituents removed from association set **D** form a candidate reactive cellular constituent set.
7914

**FIG. 79A**

7914

7916

For each cellular constituent **i** in the candidate causative cellular constituent set, determine the amount of genetic variation in the trait of interest that is explained by the eQTL of cellular constituent **i** coincident with the cQTL from the trait of interest. Rank order the cellular constituents in the candidate causative cellular constituent set based upon the amount of genetic variation in the trait of interest that is explained by each cellular constituent determined in this manner.

7918

For each eQTL of each cellular constituent **i** in the candidate causative cellular constituent set, test for the relationship:

$$P(T,Q|G) = P(T|G)P(Q|G)$$

where,

$T$ is variance in the trait of interest,

$Q$ is variance in the genome at the position where the eQTL (assoicated with cellular constituent **i**) overlaps with a cQTL linked to the trait of interest, and

$G$ is variance in the abundance level of cellular constituent **i**

7920

Optionally, determine whether each cellular constituent **i** in the candidate causative cellular constituent set includes a druggable domain

7922

Optionally, rank cellular constituents in the candidate causative cellular constituent set based on the rank assigned in step 716 and the results of step 7918 and/or step 7920

7924

Optionally, validate top ranking cellular constituents using gene knock outs/ins, transgenic construction, siRNA, drug treatments targeting candidate genes, time series experiments, *etc.*

**FIG. 79B**

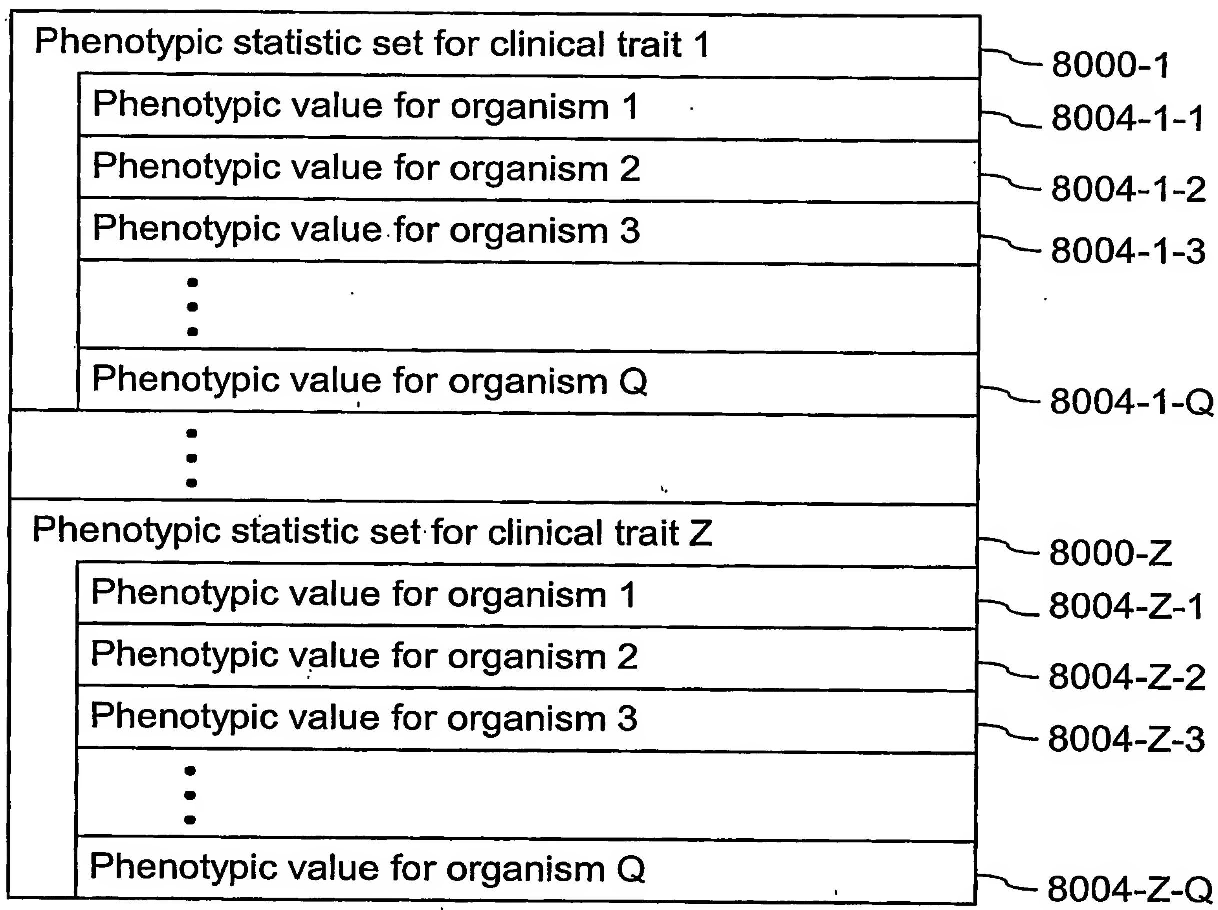

Phenotypic statistic set for clinical trait 1
Phenotypic value for organism 1
Phenotypic value for organism 2
Phenotypic value for organism 3
Phenotypic value for organism Q
Phenotypic statistic set for clinical trait Z
Phenotypic value for organism 1
Phenotypic value for organism 2
Phenotypic value for organism 3
Phenotypic value for organism Q
8000-1
8004-1-1
8004-1-2
8004-1-3
8004-1-Q
8000-Z
8004-Z-1
8004-Z-2
8004-Z-3
8004-Z-Q

FIG. 80

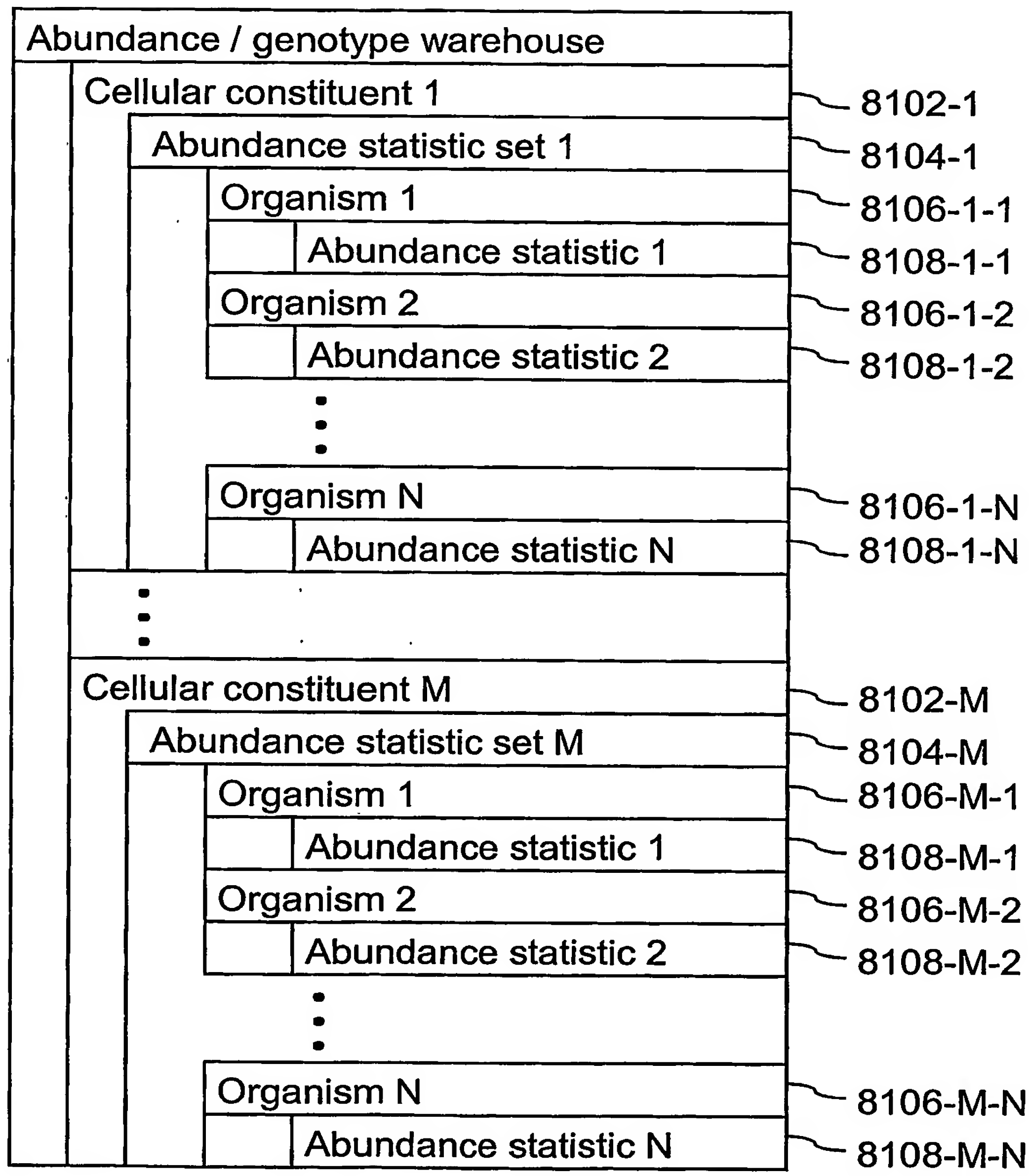

Abundance / genotype warehouse
Cellular constituent 1
Abundance statistic set 1
Organism 1
Abundance statistic 1
Organism 2
Abundance statistic 2
Organism N
Abundance statistic N
Cellular constituent M
Abundance statistic set M
Organism 1
Abundance statistic 1
Organism 2
Abundance statistic 2
Organism N
Abundance statistic N
8102-1
8104-1
8106-1-1
8108-1-1
8106-1-2
8108-1-2
8106-1-N
8108-1-N
8102-M
8104-M
8106-M-1
8108-M-1
8106-M-2
8108-M-2
8106-M-N
8108-M-N

FIG. 81

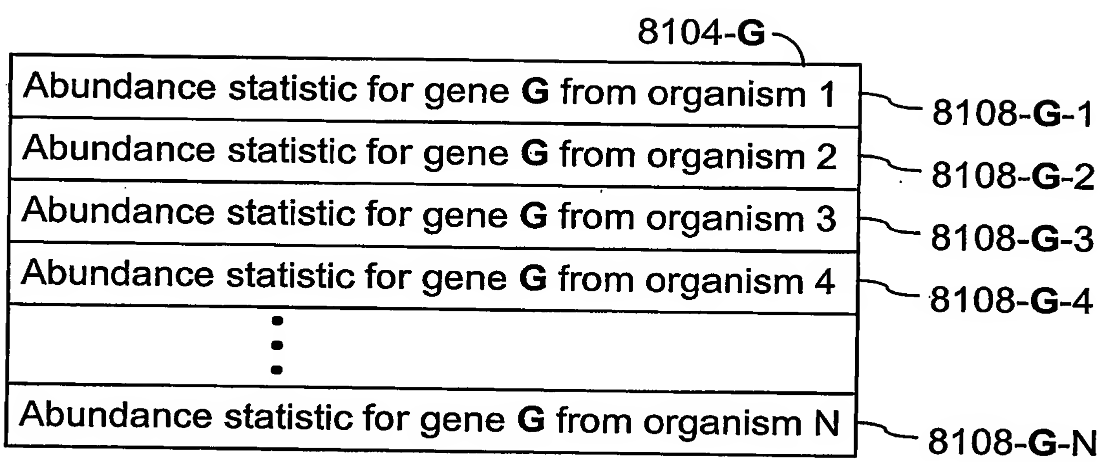

8104-G
Abundance statistic for gene G from organism 1
8108-G-1
Abundance statistic for gene G from organism 2
8108-G-2
Abundance statistic for gene G from organism 3
8108-G-3
Abundance statistic for gene G from organism 4
8108-G-4
Abundance statistic for gene G from organism N
8108-G-N

FIG. 82

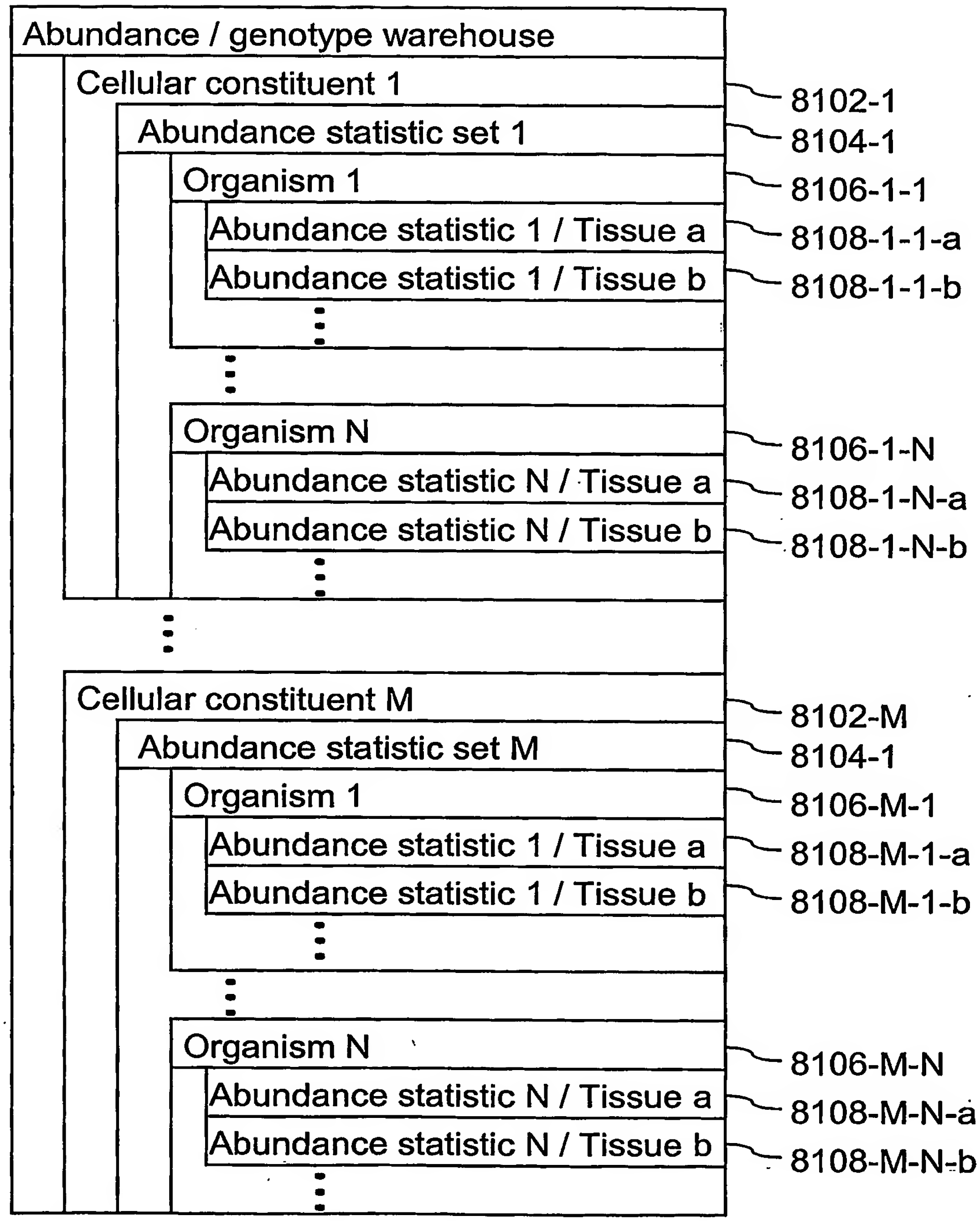
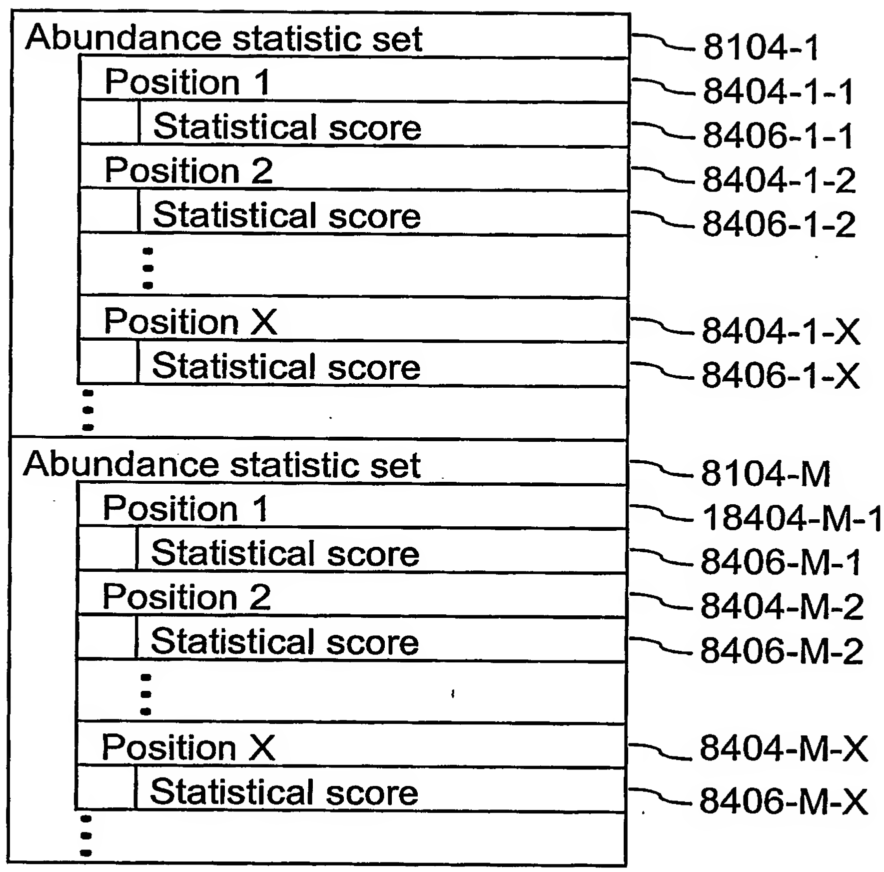

Abundance / genotype warehouse
Cellular constituent 1
Abundance statistic set 1
Organism 1
Abundance statistic 1 / Tissue a
Abundance statistic 1 / Tissue b
Organism N
Abundance statistic N / Tissue a
Abundance statistic N / Tissue b
Cellular constituent M
Abundance statistic set M
Organism 1
Abundance statistic 1 / Tissue a
Abundance statistic 1 / Tissue b
Organism N
Abundance statistic N / Tissue a
Abundance statistic N / Tissue b
8102-1
8104-1
8106-1-1
8108-1-1-a
8108-1-1-b
8106-1-N
8108-1-N-a
8108-1-N-b
8102-M
8104-1
8106-M-1
8108-M-1-a
8108-M-1-b
8106-M-N
8108-M-N-a
8108-M-N-b

FIG. 83



Abundance statistic set
Position 1
Statistical score
Position 2
Statistical score
Position X
Statistical score
Abundance statistic set
Position 1
Statistical score
Position 2
Statistical score
Position X
Statistical score
8104-1
8404-1-1
8406-1-1
8404-1-2
8406-1-2
8404-1-X
8406-1-X
8104-M
18404-M-1
8406-M-1
8404-M-2
8406-M-2
8404-M-X
8406-M-X


FIG. 84

$G \longrightarrow T$

$G \longleftarrow T$

FIG. 85A

(i) $X \longrightarrow Y \longrightarrow Z$

(ii) $X \longleftarrow Y \longrightarrow Z$

(iii) $X \longrightarrow Y \longleftarrow Z$

FIG. 85B

$Q \longrightarrow T$

FIG. 85C


FIG. 85D

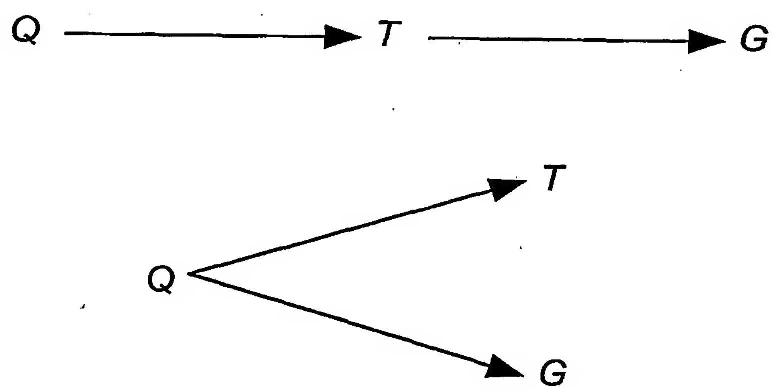



FIG. 85E

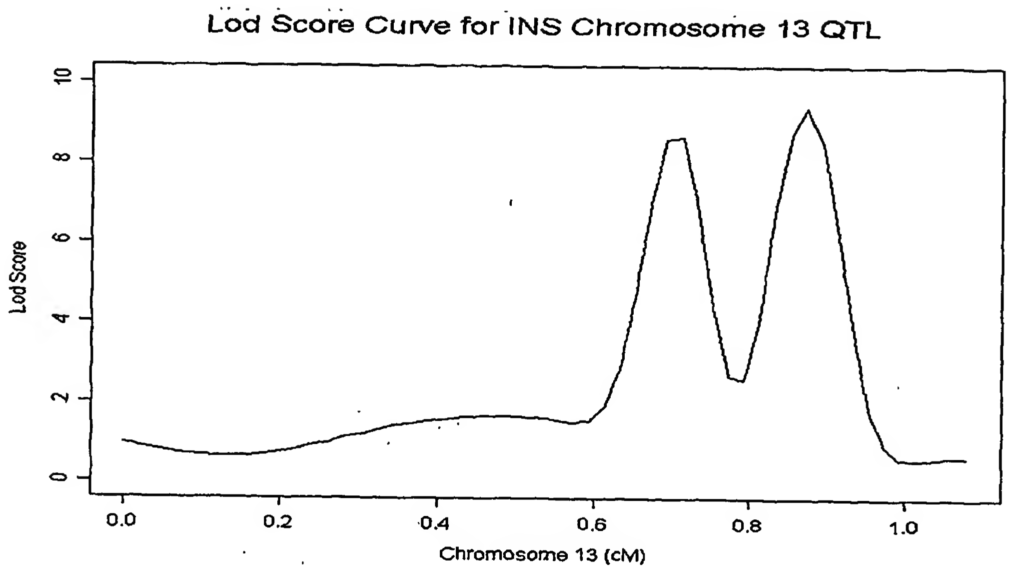



Fig. 86A

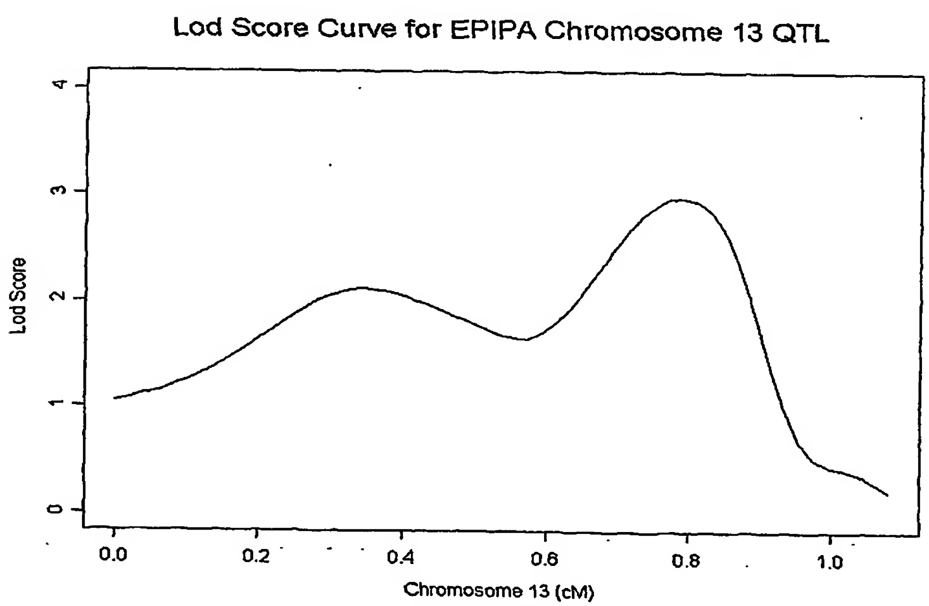



Fig. 86B

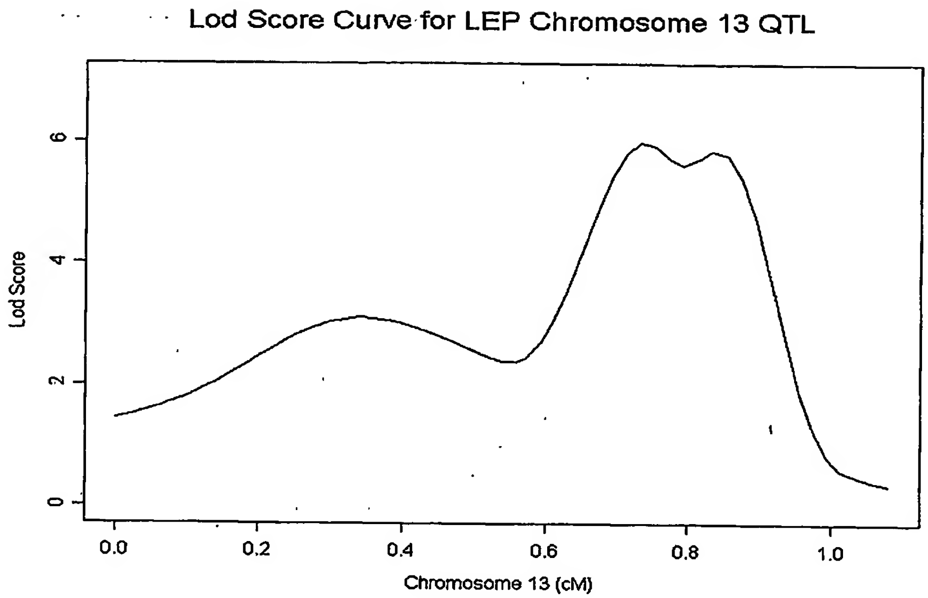



Fig. 86C

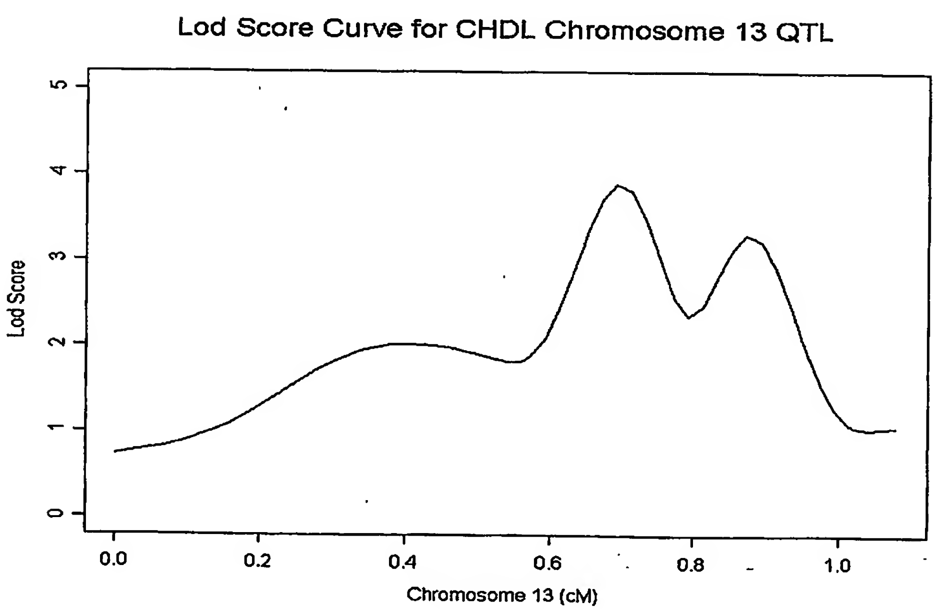



Fig. 86D

| Name | Chr Location | LOD score | site on Chr 13 (cM) |
|---|---|---|---|
| CCKar | 5 | 20.9 | 87 |
| Foxc2 | 8 | 5.2 | 85 |
| LepR | 4 | 4.1 | 110 |
| DPP IV | 2 | 3.4 | 87 |
| LIPE | 7 | 2.6 | 110 |
| CPT1a | 19 | 2.5 | 110 |
| CB2R | 4 | 2.1 | 110 |
| Orexin | 11 | 2.1 | 110 |
| CTE1 | 12 | 2 | 100 |
| RXRg | 1 | 2 | 100 |
| C330026N13Rik | 13 | 12 | 85 |
| 1810058I14Rik | 13 | 51.1 | 100 |

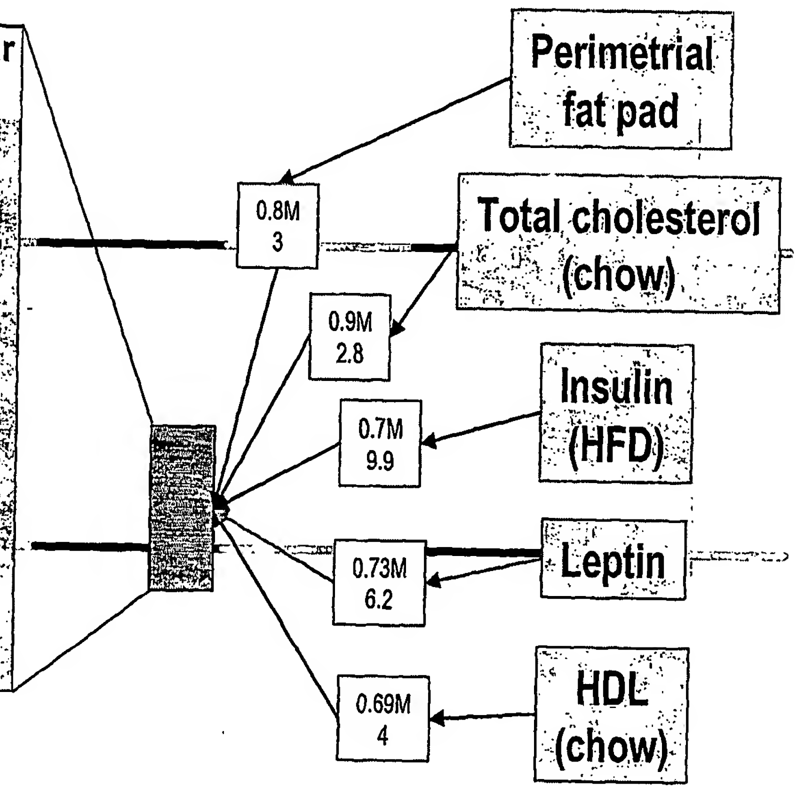
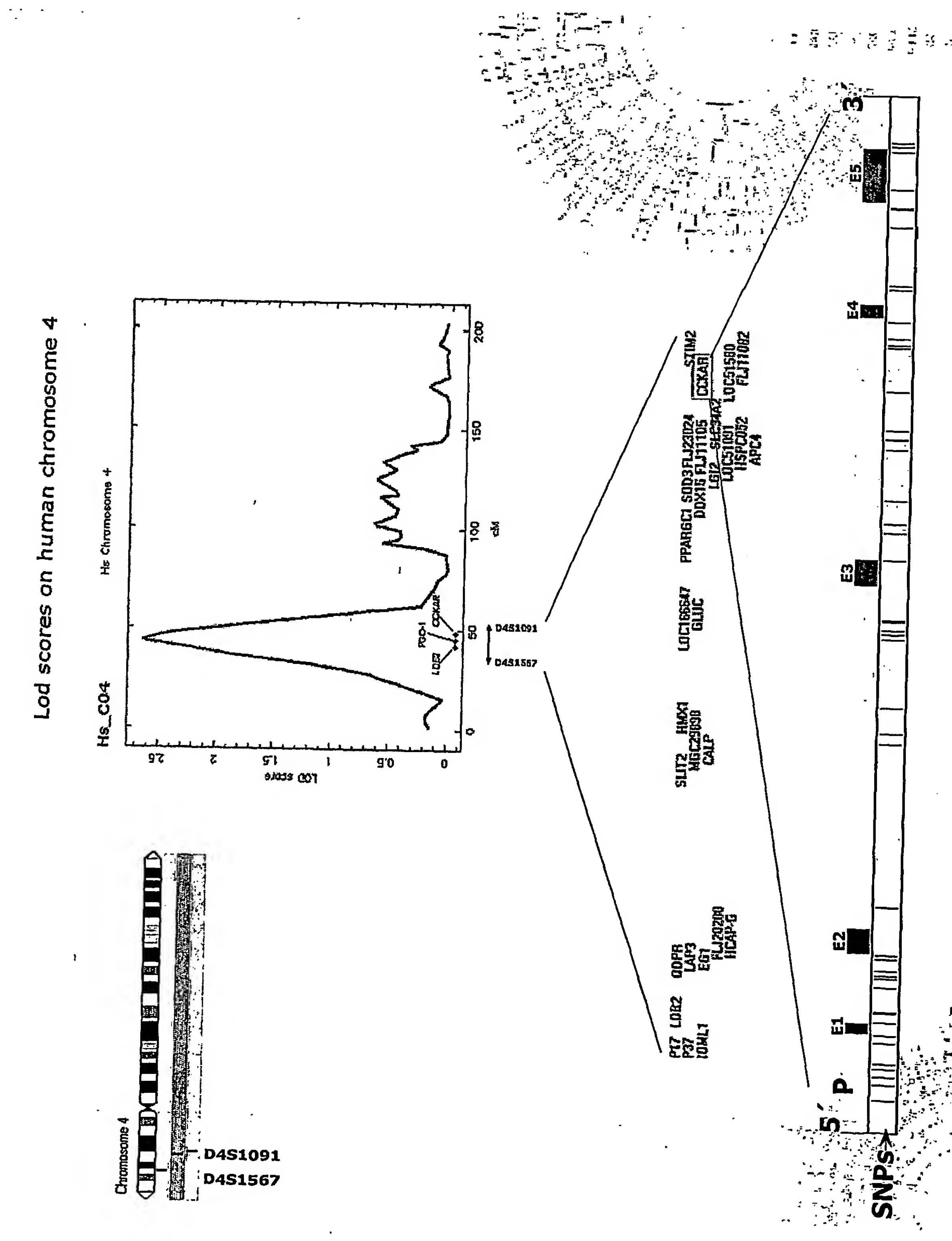



Fig. 87

1 mdvvdsllvn gsnitppcel glenetlfcl dqprpskewq pavqillysl ifllsvlgnt
61 lvitvlirnk rmrtvtnifl lslavsdlml clfcmpfnli pnllkdfifg savckttyf
121 mgtsvsvstf nlvaislery gaickplqsr vwqtkshalk viaatwclsf timtpypiys
181 nlvpftknnn qtanmcrfll pndvmqqswh tflllilfli pgivmmvayg lislelyqgi
241 kfeasqkksa kerkpsttss gkyedsdgcy lqktrpprkl elrqlstgss sranrirsns
301 saanlmakkr virmlivivv lfflcwmpif sanawraydt asaerrlsgt pisfilllsy
361 tsscvnpiiy cfmnkrfrlg fmatfpccpn pgppgargev geeeeggttg aslsrfsysh
421 msasvppq (SEQ ID NO: 30)

Fig. 88




55 SNPs identified

Fig. 89

# CCKAR haplotypes associate with high body fat in females

**Percentage body fat (top 15%) females**

| P_cor | P-unc | RRisk | PAR | Aff_frq | N_aff | Ctrl_frq | N_ctrl | Info |
|---|---|---|---|---|---|---|---|---|
| 0.002 | 8.42E-07 | 4.0 | 0.163 | 0.11 | 281 | 0.03 | 282 | 0.81 |
| 0.002 | 1.43E-06 | 4.2 | 0.163 | 0.11 | 281 | 0.03 | 279 | 0.75 |

9002

**Carrier frequency : 23% obese vs 6% thin and Relative Risk >4**

**Controls : Thin females**

Fig. 90

# CCKAR haplotypes associate with thinness in females

Thin (BMI<20) females at ages > 40 yrs

9102

| _cor | P-unc | RRisk | PAR | Aff_frq | N_aff | Ctrl_frq | N_ctrl | Info |
|---|---|---|---|---|---|---|---|---|
| 0.02 | 1.61E-05 | 4.4 | 0.119 | 0.08 | 282 | 0.02 | 421 | 0.643 |
| 0.02 | 1.84E-05 | 4.2 | 0.120 | 0.08 | 282 | 0.02 | 421 | 0.647 |

**Carrier frequency : 17% thin vs 4% obese and Relative Risk >4**

**Controls : Obese females**

Fig. 91